Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 1–12

Approach Merotopies and Associated Near Sets

James Peters[a,b,*], Surabhi Tiwari[c], Rashmi Singh[d]

[a]*Computational Intelligence Laboratory, University of Manitoba, Winnipeg, Manitoba R3T 5V6, Canada.*
[b]*School of Mathematics & Computer / Information Sciences, University of Hyderabad, Central Univ. P.O., Hyderabad 500046, India.*
[c]*Department of Mathematics, Motilal Nehru National Institute of Technology, Allahabad- 211 004, U.P., India.*
[d]*Department of Mathematics, Amity Institute of Applied Sciences, Noida, U.P., India.*

Abstract

This article introduces associated near sets of a collection of sets. The proposed approach introduces a means of defining as well as describing an ε-approach merotopy in terms of the members of associated sets of collections that are sufficiently near. A characterization for continuous functions is established using associated near sets. This article also introduces p-containment considered in the context of near sets. An application of the proposed approach is given in terms of digital image classification.

Keywords: Approach space, associated set, p-containment, merotopy, near sets.
2010 MSC: Primary 26A21, Secondary 26A24, 54D35, 54A20, 54E99, 18B30.

1. Introduction

For any real-valued function f of a real variable, the associated sets of f (Agronsky, 1982) are the sets

$$E^{\alpha}(f) = \{x : f(x) < \alpha\} \text{ and } E_{\alpha}(f) = \{x : f(x) > \alpha\},$$

where $\alpha \in \mathbb{R}$ (the set of all real numbers). Many classes of functions can be characterized in terms of their associated sets. The study of associated sets of a function started in 1922 (Coble, 1922) and elaborated in (Zahorski, 1950; Bruckner, 1967; Agronsky, 1982; Petrakiev, 2009). For example, a function is continuous, if and only if, all of its associated sets are open, a function is approximately continuous if, and only if, all of its associated sets F sets with the property that every point of an associated set is a point of Lebesgue density of that set. More generally, A. Bruckner (Bruckner, 1967, p. 228) has shown that if κ is a class of functions characterized in terms of an associated set P and h is a homeomorphism, then the associated sets of the function $h \circ f$ are all members

*Corresponding author
Email addresses: jfpeters@ee.umanitoba.ca (James Peters), au.surabhi@gmail.com (Surabhi Tiwari)

of P and $h \circ f \in \kappa$. S. Agronsky (Agronsky, 1982, p. 767) has observed that an associated set for a function in $\mathcal{M}_i$ must be 'more dense' near each of its members than an associated set for a function in $\mathcal{M}_{i-1}$.

In this paper, associated sets defined in terms of ε-approach merotopies are considered. In particular, we consider associated sets containing members that are sufficiently near each other relative to ε-approach merotopies. Carrying forward the idea of defining and characterizing a function in terms of an associated set, it is possible to define and characterize an approach merotopy in terms of an associated set of collections. Using the concept of associated sets, an equivalent condition for continuous functions is obtained.

2. Preliminaries

Let X be a nonempty ordinary set. The power set of X is denoted by $\mathcal{P}(X)$, the family of all collections of subsets of $\mathcal{P}(X)$ is denoted by $\mathcal{P}^2(X)$. We denote by $\aleph_0$ the first infinite cardinal number, by J an arbitrary index set, and $|A|$ is the cardinality of A, where $A \subseteq X$. For $\mathcal{A}, \mathcal{B} \in \mathcal{P}^2(X)$, we say $\mathcal{A} \vee \mathcal{B} \equiv \{A \cup B : A \in \mathcal{A}, B \in \mathcal{B}\}$; $\mathcal{A}$ *corefines* $\mathcal{B}$ (written as $\mathcal{A} \prec \mathcal{B}$), if and only if, for all $A \in \mathcal{A}$, there exists $B \in \mathcal{B}$ such that $B \subseteq A$. For $\mathcal{A} \subseteq \mathcal{P}(X)$, $stack(\mathcal{A}) = \{A \subseteq X : B \subseteq A$, for some $B \in \mathcal{A}\}$ and $sec(\mathcal{A}) = \{B \subseteq X : A \cap B \neq \emptyset$, for all $A \in \mathcal{A}\} = \{B \subseteq X : X - B \notin$ stack$(\mathcal{A})\}$. Observe that $\text{sec}^2(\mathcal{A}) = \text{stack}(\mathcal{A})$, for all $\mathcal{A} \in \mathcal{P}^2(X)$. A *filter* on X is a nonempty subset $\mathcal{F}$ of $\mathcal{P}(X)$ satisfying: $\emptyset \notin \mathcal{F}$; if $A \in \mathcal{F}$ and $A \subseteq B$, then $B \in \mathcal{F}$; and if $A \in \mathcal{F}$ and $B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$. A maximal filter on X is called an *ultrafilter* on X. A *grill* on X is a subset $\mathcal{G}$ of $\mathcal{P}(X)$ satisfying: $\emptyset \notin \mathcal{G}$; if $A \in \mathcal{G}$ and $A \subseteq B$, then $B \in \mathcal{G}$; and if $A \cup B \in \mathcal{G}$, then $A \in \mathcal{G}$ or $B \in \mathcal{G}$. Note that for any $x \in X, \dot{x} = \{A \subseteq X : x \in A\}$ is an ultrafilter on X, which is also a grill on X. There is one-to-one correspondence between the set of all filters and the set of all grills on X by the relation: $\mathcal{F}$ is a filter on X if and only if $\text{sec}(\mathcal{F})$ is a grill on X; and $\mathcal{G}$ is a grill on X if and only if, $\text{sec}(\mathcal{G})$ is a filter on X.

In its most basic form, an approach merotopy is a measure of the nearness of members of a collection. For collections $\mathcal{A}, \mathcal{B} \in \mathcal{P}^2(X)$, a function $\nu : \mathcal{P}^2(X) \times \mathcal{P}^2(X) :\longrightarrow [0, \infty]$ satisfying a number of properties is a called an ε-approach merotopy. A pair of collections are near, provided $\nu(\mathcal{A}, \mathcal{B}) = 0$. For $\varepsilon \in (0, \infty]$, the pair $\mathcal{A}, \mathcal{B}$ are *sufficiently near*, provided $\nu(\mathcal{A}, \mathcal{B}) < \varepsilon$.

Let cl be a Kuratowski closure operator on X. Then the topological space (X, cl) is called a *symmetric topological space* if and only if $x \in cl(\{y\}) \Longrightarrow y \in cl(\{x\})$, for all $x, y \in X$.

Definition 2.1. A function $\delta : X \times \mathcal{P}(X) \longrightarrow [0, \infty]$ is called a distance on X (Lowen, 1997; Lowen *et al.*, 2003) if for any $A, B \subseteq X$ and $x \in X$, the following conditions are satisfied:
(**D**.1) $\delta(x, \{x\}) = 0$,
(**D**.2) $\delta(x, \emptyset) = \infty$,
(**D**.3) $\delta(x, A \cup B) = \min\{\delta(x, A), \delta(x, B)\}$,
(**D**.4) $\delta(x, A) \leq \delta(x, A^{(\alpha)}) + \alpha$, for all $\alpha \in [0, \infty]$, where $A^{(\alpha)} \doteqdot \{x \in X : \delta(x, A) \leq \alpha\}$.
The pair (X, δ) is called an approach space.

Definition 2.2. A generalized approach space (X, ρ) (Peters & Tiwari, 2011, 2012) is a nonempty set X equipped with a generalized distance function $\rho : \mathcal{P}(X) \times \mathcal{P}(X) \longrightarrow [0, \infty]$, if and only if, for all nonempty subsets $A, B, C \in \mathcal{P}(X)$, ρ satisfies properties (A.1)-(A.5), *i.e.*,

(**A**.1) $\rho(A, A) = 0$,
(**A**.2) $\rho(A, \emptyset) = \infty$,
(**A**.3) $\rho(A, B \cup C) = \min\{\rho(A, B), \rho(A, C)\}$,
(**A**.4) $\rho(A, B) = \rho(B, A)$,
(**A**.5) $\rho(A, B) \leq \rho(A, B^{(\alpha)}) + \alpha$, for every $\alpha \in [0, \infty]$, where $B^{(\alpha)} = \{x \in X : \rho(\{x\}, B) \leq \alpha\}$.

It has been observed that the notion of distance in an approach space is closely related to the notion of nearness (Khare & Tiwari, 2012, 2010; Tiwari, Jan. 2010). In particular, consider the Čech distance between sets.

Definition 2.3. Čech Distance (Čech, 1966). For nonempty subsets $A, B \in \mathcal{P}(X)$, $\rho(a, b)$ is the standard distance between $a \in A, b \in B$ and the Čech distance $D_\rho : \mathcal{P}(X) \times \mathcal{P}(X) \longrightarrow [0, \infty]$ is defined by

$$D_\rho(A, B) \doteqdot \begin{cases} \inf\{\rho(a, b) : a \in A, b \in B\}, & \text{if } A \text{ and } B \text{ are not empty,} \\ \infty, & \text{if } A \text{ or } B \text{ is empty.} \end{cases}$$

Remark. Observe that (X, D_ρ) is a generalized approach space. The distance $D_\rho(A, B)$ is a variation of the distance function introduced by E. Čech in his 1936–1939 seminar on topology (Čech, 1966) (see, also, (Beer *et al.*, 1992; Hausdorff, 1914*a*; Leader, 1959)).

3. Approach merotopic spaces

Definition 3.1. Let $\varepsilon \in (0, \infty]$. Then a function $\nu : \mathcal{P}^2(X) \times \mathcal{P}^2(X) \longrightarrow [0, \infty]$ is an ε-approach merotopy on X, if and only if, for any collections $\mathcal{A}, \mathcal{B}, C \in \mathcal{P}^2(X)$, the properties (AN.1)-(AN.5) are satisfied.

(AN.1) $\mathcal{A} \prec \mathcal{B} \Longrightarrow \nu(C, \mathcal{A}) \leq \nu(C, \mathcal{B})$,

(AN.2) $\mathcal{A} \neq \emptyset, \mathcal{B} \neq \emptyset$ and $(\bigcap \mathcal{A}) \cap (\bigcap \mathcal{B}) \neq \emptyset \Longrightarrow \nu(\mathcal{A}, \mathcal{B}) < \varepsilon$,

(AN.3) $\nu(\mathcal{A}, \mathcal{B}) = \nu(\mathcal{B}, \mathcal{A})$ and $\nu(\mathcal{A}, \mathcal{A}) = 0$,

(AN.4) $\mathcal{A} \neq \emptyset \Longrightarrow \nu(\emptyset, \mathcal{A}) = \infty$,

(AN.5) $\nu(C, \mathcal{A} \vee \mathcal{B}) \geq \nu(C, \mathcal{A}) \wedge \nu(C, \mathcal{B})$.

The pair (X, ν) is termed as an *ε-approach merotopic space*.

For an ε-approach merotopic space (X, ν), we define: $cl_\nu(A) \doteqdot \{x \in X : \nu(\{\{x\}\}, \{A\}) < \varepsilon\}$, for all $A \subseteq X$. Then cl_ν is a Čech closure operator on X.

Let $cl_\nu(\mathcal{A}) \doteqdot \{cl_\nu(A) : A \in \mathcal{A}\}$. Then an ε-approach merotopy ν on X is called an *ε-approach nearness* on X, if the following condition is satisfied:

(AN.6) $\nu(cl_\nu(\mathcal{A}), cl_\nu(\mathcal{B})) \geq \nu(\mathcal{A}, \mathcal{B})$.

In this case, cl_ν is a Kuratowski closure operator on X.

Lemma 3.1. *Let $\varepsilon \in (0,\infty]$, and let (X,ν) and (Y,ν') be ε-approach nearness spaces. Then $f : (X,\nu) \longrightarrow (Y,\nu')$ is a contraction if and only if*
$\nu(f^{-1}(\mathcal{A}), f^{-1}(\mathcal{B})) \geq \nu'(\mathcal{A},\mathcal{B})$, *for all* $\mathcal{A},\mathcal{B} \in \mathcal{P}^2(Y)$.

Example 3.1. Let D_ρ be a gap functional. Then the function $\nu_{D_\rho} : \mathcal{P}^2(X) \times \mathcal{P}^2(X) \longrightarrow [0,\infty]$ defined as

$$\nu_{D_\rho}(\mathcal{A},\mathcal{B}) \doteqdot \sup_{A\in\mathcal{A},B\in\mathcal{B}} D_\rho(A,B); \quad \nu_{D_\rho}(\mathcal{A},\mathcal{A}) \doteqdot \sup_{A\in\mathcal{A}} D_\rho(A,A) = 0,$$

is an ε-approach merotopy on X. Define $cl_\rho(A) = \{x \in X : \rho(\{x\},A) < \varepsilon\}$, $A \subseteq X$. Then cl_ρ is a Čech closure operator on X. Further, if $\rho(cl_\rho(A), cl_\rho(B)) \geq \rho(A,B)$, for all $A,B \subseteq X$, then cl_ρ is a Kuratowski closure operator on X, and we call ρ as an *ε-approach function* on X; and (X,ρ) is an *ε-approach space*. In this case, ν_{D_ρ} is an ε-approach nearness on X.

So, there are many instances of ε-approach nearness on X just as there are many instances of ε-approach spaces (Lowen, 1997) and metric spaces on X.

Definition 3.2. Near and Almost Near Collections
For collections $\mathcal{A},\mathcal{B} \in \mathcal{P}^2(X)$, assume that the function $\nu : \mathcal{P}^2(X) \times \mathcal{P}^2(X) :\longrightarrow [0,\infty]$ is an ε-approach merotopy. A pair of collections are *near*, provided $\nu(\mathcal{A},\mathcal{B}) = 0$. For $\varepsilon \in (0,\infty]$, the pair $\mathcal{A},\mathcal{B}$ are *ε-near* (*almost near*), provided $\nu(\mathcal{A},\mathcal{B}) < \varepsilon$ (Peters & Tiwari, 2011). Otherwise, collections $\mathcal{A},\mathcal{B}$ are far, *i.e.*, *sufficiently apart*, provided $\nu(\mathcal{A},\mathcal{B}) \geq \varepsilon$.

4. Associated collections

It is possible to characterise ε-approach merotopies in terms of associated collections.

Definition 4.1. Associated Collections of an ε-Approach Merotopy
Let X denote an ordinary nonempty set and let $\mathcal{A} \in \mathcal{P}^2(X)$ denote collections of subsets of X. Suppose that $\varepsilon \in (0,\infty]$ and ν be an ε-approach merotopic space. The upper associated set of $\mathcal{A}$ with respect to ν is defined by

$$E^\varepsilon(\mathcal{A}) \doteqdot \{\mathcal{B} \in \mathcal{P}^2(X) : \ \nu(\mathcal{A},\mathcal{B}) > \varepsilon\}.$$

and the lower associated set of $\mathcal{A}$ with respect to ν is defined by

$$E_\varepsilon(\mathcal{A}) \doteqdot \{\mathcal{B} \in \mathcal{P}^2(X) : \ \nu(\mathcal{A},\mathcal{B}) < \varepsilon\}.$$

Example 4.1. Let D_ρ be a gap functional. For $\mathcal{A},\mathcal{B} \in \mathcal{P}^2(X)$, the function $\nu_{D_\rho} : \mathcal{P}^2(X) \times \mathcal{P}^2(X) \longrightarrow [0,\infty]$ is defined by

$$\nu_{D_\rho}(\mathcal{A},\mathcal{B}) \doteqdot \sup_{A\in\mathcal{A},B\in\mathcal{B}} D_\rho(A,B); \quad \nu_{D_\rho}(\mathcal{A},\mathcal{A}) \doteqdot \sup_{A\in\mathcal{A}} D_\rho(A,A) = 0.$$

From Def. 4.1, $E_\varepsilon(\mathcal{A})$ is the lower associated set of $\mathcal{A}$ for a given $\varepsilon \in \mathbb{R}$. Similarly, obtain the upper associated set $E^\varepsilon(\mathcal{A})$ of $\mathcal{A}$ as a collection $\mathcal{B} \in \mathcal{P}^2(X)$, provided $\nu_{D_\rho}(\mathcal{A},\mathcal{B}) > \varepsilon$.

Additional examples of lower and upper associated collections are given next.

Example 4.2. Let (X, ν) be an ε-approach nearness on X, $r < \varepsilon < \infty$ and $\varepsilon' < \varepsilon$. Then

(**ASet**.1) Associated sets $E^{\varepsilon}(\mathcal{A}), E_{\varepsilon}(\mathcal{A})$ of $\mathcal{A}$ with respect to $\nu_1 : \mathcal{P}^2(X) \times \mathcal{P}^2(X) \longrightarrow [0, \infty]$ such that

$$\nu_1(\mathcal{A}, \mathcal{B}) = \begin{cases} \infty, & \text{if } \emptyset \in \mathcal{A} \text{ or } \emptyset \in \mathcal{B}, \\ r, & \text{otherwise,} \end{cases}$$

is defined by:
if $\emptyset \in \mathcal{A}$, $E^{\varepsilon}(\mathcal{A}) = \mathcal{P}^2(X)$ and $E_{\varepsilon}(\mathcal{A}) = \emptyset$,
if $\emptyset \notin \mathcal{A}$, $E^{\varepsilon}(\mathcal{A}) = \{\mathcal{A} \in \mathcal{P}^2(X) : \emptyset \in \mathcal{B}\}$ and $E_{\varepsilon}(\mathcal{A}) = \{\mathcal{A} \in \mathcal{P}^2(X) : \emptyset \notin \mathcal{B}\}$.

(**ASet**.2) Associated sets $E^{\varepsilon}(\mathcal{A}), E_{\varepsilon}(\mathcal{A})$ of $\mathcal{A}$ with respect to $\nu_2 : \mathcal{P}^2(X) \times \mathcal{P}^2(X) \longrightarrow [0, \infty]$ such that

$$\nu_2(\mathcal{A}, \mathcal{B}) = \begin{cases} \infty, & \text{if } \emptyset \in \mathcal{A} \text{ or } \emptyset \in \mathcal{B}, \\ \inf\{\nu(\mathcal{A}, \mathcal{B}), \varepsilon'\}, & \text{otherwise,} \end{cases}$$

is defined by:
if $\emptyset \in \mathcal{A}$, $E^{\varepsilon}(\mathcal{A}) = \mathcal{P}^2(X)$ and $E_{\varepsilon}(\mathcal{A}) = \emptyset$,
if $\emptyset \notin \mathcal{A}$, $E^{\varepsilon}(\mathcal{A}) = \{\mathcal{A} \in \mathcal{P}^2(X) : \emptyset \in \mathcal{B}\}$ and $E_{\varepsilon}(\mathcal{A}) = \{\mathcal{A} \in \mathcal{P}^2(X) : \emptyset \notin \mathcal{B}\}$.

Proposition 1. *A collection in the lower associated set of $\mathcal{A}$ with respect to the ε-approach merotopy ν is sufficiently near $\mathcal{A}$.*

Proof. Assume $\mathcal{B} \in E_{\varepsilon}(\mathcal{A})$, the lower associated set of $\mathcal{A}$ with respect to ν. From Def. 3.2, $\mathcal{A}, \mathcal{B}$ are sufficiently near. □

Proposition 2. *A collection in upper associated set of $\mathcal{A}$ with respect to the ε-approach merotopy ν are sufficiently apart.*

Proof. Immediate from from Def. 4.1 and Def. 3.2. □

We now present a characterization for continouous functions.

Theorem 4.1. *Let ν_X and ν_Y be ε-approach merotopies on X and Y, respectively. A mapping $f : X \longrightarrow Y$ is continuous, if and only if, $\mathcal{A} \in E_{\varepsilon}(x) \implies f(\mathcal{A}) \in E_{\varepsilon}(f(x))$, for all $\mathcal{A} \in \mathcal{P}^2(X)$ and for all $x \in X$.*

Proof. Let $f : X \longrightarrow Y$ be continuous, $x \in X$ and $\mathcal{A} \in \mathcal{P}^2(X)$. Suppose that $\mathcal{A} \in E_{\varepsilon}(x)$. Then $\nu(\mathcal{A}, \{\{x\}\}) < \varepsilon$, which gives $\nu(\{A\}, \{\{x\}\}) < \varepsilon$, for all $A \in \mathcal{A}$. That is, $x \in cl_{\nu_X}(A)$, for all $A \in \mathcal{A}$. Consequently, $f(x) \in f(cl_{\nu_X}(A)) \subseteq cl_{\nu_Y}(f(A))$, for all $A \in \mathcal{A}$. Hence, $f(\mathcal{A}) \in E_{\varepsilon}(f(x))$. The converse is obvious. □

Definition 4.2. Finite Strong Containment Property (Agronsky, 1982).
Let p be a property defined for sets of real numbers with respect to sets containing them. If $A \subset B$, then A is p-contained in B (written $A \underset{p}{\subset} B$), provided A has the property p with respect to B. Put $k \in [0, \infty)$. Then p is a finite strong containment property, provided

(**p**.1) If $A \underset{p}{\subset} B \subset F$ and p is defined for $A \subset F$, then $A \underset{p}{\subset} F$,

(**p**.2) If $A \subset B \underset{p}{\subset} F$, then $A \underset{p}{\subset} F$,

(**p**.3) If, for each $n \in \mathbb{N}, E_n \underset{p}{\subset} F_n$, then $\bigcup_{n=1}^{k} E_n \underset{p}{\subset} \bigcup_{n=1}^{k} F_n$.

Example 4.3. Strong Containment of Sufficiently Near Collections
Put $\varepsilon \in (0, \infty]$. Let (X, ν) be an ε-approach nearness on X and $p \div$ 'sufficiently near' defined for $\mathcal{A}, \mathcal{B} \in \mathcal{P}(X)$ such that $\nu(\mathcal{A}, \mathcal{B}) < \varepsilon$. From Example 4.2, assume $\mathcal{A}, \mathcal{B} \in E_\varepsilon(\nu_2)$ and $\mathcal{A} \subset \mathcal{B}$, then $\mathcal{A} \underset{p}{\subset} \mathcal{B}$.

Proof.

(**p**.1) Assume $\mathcal{A}, \mathcal{B}, \mathcal{C} \in E_\varepsilon(\nu_2)$. By definition, $\mathcal{A} \underset{p}{\subset} \mathcal{B}$. Assume $\mathcal{B} \subset \mathcal{C}$, then $\mathcal{A} \underset{p}{\subset} \mathcal{B} \subset \mathcal{C}$. Since $\mathcal{B}, \mathcal{C} \in E_\varepsilon(\nu_2)$, then $\mathcal{A} \underset{p}{\subset} \mathcal{C}$.

(**p**.2) Assume $\mathcal{A}, \mathcal{B}, \mathcal{C} \in E_\varepsilon(\nu_2)$ and that $\mathcal{A} \subset \mathcal{B} \subset \mathcal{C}$. By definition, $\mathcal{A} \subset \mathcal{B} \underset{p}{\subset} \mathcal{C}$ and by assumption $\mathcal{A} \subset \mathcal{C}$. Since $\mathcal{A}, \mathcal{C} \in E_\varepsilon(\nu_2)$, then $\mathcal{A} \underset{p}{\subset} \mathcal{C}$.

(**p**.3) The proof of this strong containment property follows by mathematical induction.

□

5. Description-based neighbourhoods

For N. Bourbaki, a set is a neighbourhood of each of its points if, and only if, the set is open (Bourbaki, 1971, §1.2) (Bourbaki, 1966, §1.2, p. 18). A set A is *open*, if and only if, for each $x \in A$, all points *sufficiently near*[1] x belong to A.

For a Hausdorff neighbourhood (denoted by N_r), *sufficiently near* is explained in terms of the distance between points y and x being less than some radius r (Hausdorff, 1914*b*, §22). In other words, a Hausdorff neighbourhood of a point is an open set such that each of its points is sufficiently close to its centre.

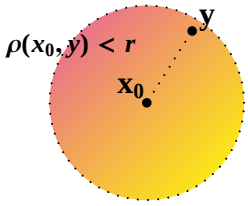


Figure 1: Nbd $N_r(x_0)$

Traditionally, nearness of points is measured in terms of the location of the points. Let $\rho : X \times X :\to [0, \infty]$ denote the standard distance[2] between points in X. For $r \in (0, \infty]$, a neighbourhood of $x_0 \in X$ is the set of all $y \in X$ such that $\rho(x_0, y) < r$ (see, *e.g.*, Fig. 1, where the distance $\rho(x, y)$ between each pair x_0, y is less than r in the neighbourhood). In that case, a neighbourhood is called an open ball (Engelking, 1989, §4.1) or spherical neighbourhood (Hocking & Young, 1988, §1-4). In the plane, the points in a spherical neighhourhood (nbd) are contained in the interior of a circle.

Next, an alternative to a spherical neighbourhood is called a visual neighbourhood (denoted nbd_v), which stems from recent work on descriptively near sets (Naimpally & Peters, 2013; Peters, 2013; Peters & Naimpally, 2012).

[1] *...tous les points assez voisins d'un point x* (Bourbaki, 1971, p. TG I.3)
[2] *i.e.*, for $x, y \in X \subset, \rho(x, y) = |x - y|$.

Definition 5.1. Visual Neighbourhood
A visual nbd_v of a point x_0 (denoted N_{r_ϕ}) is an open set A such that the visual information values extracted from all of the points in A are sufficiently near the corresponding visual information values at x_0. Let ϕ denote a probe function used to extract visual information from a point in nbd_v. Sufficient nearness of points in a visual nbd_v is defined in terms of bound r_ϕ, a real number. That is, points $x_0, x \in A$ are sufficiently near, *i.e.*, provided

$$\rho_\phi(x_0, y) = |\phi(x_0) - \phi(y)| < r_\phi.$$

Example 5.1. Visual Neighbourhood in a Drawing
In its simplest form (see, *e.g.*, Fig. 2), a nbd_v (denoted by N_{r_ϕ}) is defined in terms of a real-valued probe function ϕ used to extract visual information from the pixels in a digital image, reference point x_0 (*not* necessarily the centre of the nbd_v) and 'radius' r_ϕ such that

$X =$ {drawing visual pixels}, $x, y \in X$,

$\phi : X \to [0, \infty]$, (probe function, *e.g.*, probe $\phi(x)$ = pixel x intensity),

$\rho_\phi(x_0, y) = |\phi(x_0) - \phi(y)|$, (visual distance),

$x_0 \in X$,(nbd_v reference point),

$r_\phi \in (0, \infty]$, (sufficient nearness bound),

$N_{r_\phi}(x_0) = \{y \in X : \rho_\phi(x_0, y) < r_\phi\}$, (visual nbd_v).

At this point, observe that the appearance of a visual neighbourhood can be quite different from the appearance of a spherical neighbourhood. For this reason, x_0 is called a *reference point* (not a centre) in a nbd_v. A visual neighbourhood results from a consideration of the features of a point in the neighbourhood and the measurement of the distance between neighbourhood points[3]. For example, $\phi(x_0)$ in Fig. 2 is a description of x_0 (probe ϕ is used to extract a feature value from x in the form of pixel intensity). Usually, a complete description of a point x in a nbd_v is in the form of a feature vector containing probe function values extracted from x (see, *e.g.*, (Henry, 2010, §4), for a detailed explanation of the near set approach to perceptual object description). Observe that the members $y \in N_{r_\phi}(x_0)$ in the visual neighbourhood in Fig. 2 have descriptions that are *sufficiently near* the description of the reference point x_0.

For example, each of the points in the green shaded regions in Fig. 2 have intensities that are very close to the intensity of the point x_0. By contrast, many points in the purple shaded region have higher intensities (*i.e.*, more light) than the pixel at x_0, For example, consider the intensities of the points in the visual nbd represented by the green wedge-shaped region and some outlying green circular regions and the point x_4 in the purple region in Fig. 2, where

$r_\phi = 5$ low intensity difference,

[3]It is easy to prove that *a visual neighbourhood is an open set*

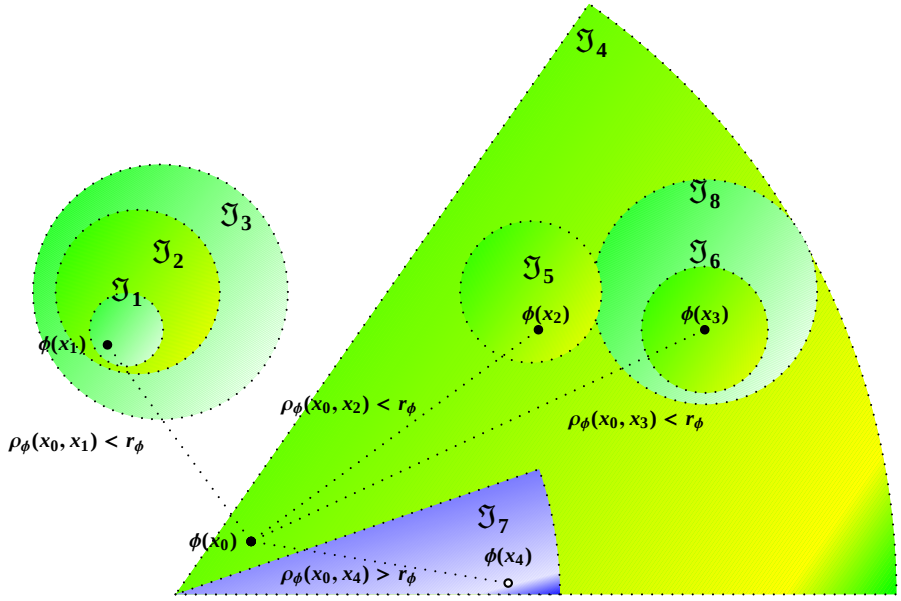


Figure 2: Sample Visual Nbd $N_{r_\phi}(x_0)$ in a Drawing

$\rho_\phi(x_0, x_1) = |\phi(x_0) - \phi(x_1)| < r_\phi$,

$\rho_\phi(x_0, x_2) = |\phi(x_0) - \phi(x_2)| < r_\phi$,

$\rho_\phi(x_0, x_3) = |\phi(x_0) - \phi(x_3)| < r_\phi$, but

$\rho_\phi(x_0, x_4) = |\phi(x_0) - \phi(x_4)| > r_\phi$, where $\phi(x_4)$ = high intensity (white).

In the case of the point x_4 in Fig. 2, the intensity is high (close to white), *i.e.*, $\phi(x_4) \sim 255$. By contrast the point x_0 has low intensity (less light), *e.g.*, $\phi(x_0) \sim 20$. Assume $r_\phi = 5$. Hence, $|\phi(x_0) - \phi(x_4)| > r_\phi$. As in the case of C. Monet's paintings[4], the distance between probe function values representing visual information extracted from image pixels can be sufficiently near a centre x_0 (perceptually) but the pixels themselves can be *far apart*, *i.e.*, not sufficiently near, if one considers the locations of the pixels.

Remark. **Filters and Grills**
In Fig. 2, observe that $\mathcal{F}_1 = \mathfrak{I}_1 \subset \mathfrak{I}_2 \subset \mathfrak{I}_3$ is a filter. Again, observe that $\mathcal{F}_2 = \{\mathfrak{I}_4, \mathfrak{I}_6, \mathfrak{I}_8\}$ is a filter. It can be shown that the set $\mathcal{G} = \{\mathfrak{I}_4, \mathfrak{I}_6, \mathfrak{I}_8\}$ is a grill.

Proof. Let $A = \mathfrak{I}_6, B = \mathfrak{I}_4$ in Fig. 2. From $\mathcal{F}_2$, we know that $\mathfrak{I}_6 \subset \mathfrak{I}_4$ and $\mathfrak{I}_4 \subset \mathcal{G}$. Then $B \in \mathcal{G}$. Observe that $\mathfrak{I}_5 \cup \mathfrak{I}_6 \in \mathcal{G}$, then $\mathfrak{I}_5 \in \mathcal{G}$ or $\mathfrak{I}_6 \in \mathcal{G}$.

[4] A comparison between Z. Pawlak's and C. Monet's waterscapes is given in Peters (2011).

In addition, let X denote the set of regions shown in Fig. 2. Obviously, $\mathcal{F} = \{\mathfrak{I}_4, \mathfrak{I}_6, \mathfrak{I}_8\}$ is a filter, if and only if, $sec(\mathcal{F})$ is a grill $\mathcal{G}_2$ on X. Further, $\mathcal{G}_3 = \{\mathfrak{I}_1, \mathfrak{I}_2, \mathfrak{I}_3\}$ is a grill, if and only if, $sec(\mathcal{G}_3)$ is a filter. □

Example 5.2. Sample Associated Sets
Let D_{ρ_ϕ} be a gap functional such that

$$D_{\rho_\phi}(A, B) \div \begin{cases} \inf\{\rho_\phi(a, b) : a \in A, b \in B\}, & \text{if } A \text{ and } B \text{ are not empty,} \\ \infty, & \text{if } A \text{ or } B \text{ is empty.} \end{cases}$$

Then the function $\nu_{D_{\rho_\phi}} : \mathcal{P}^2(X) \times \mathcal{P}^2(X) \longrightarrow [0, \infty]$ defined as

$$\nu_{D_{\rho_\phi}}(\mathcal{A}, \mathcal{B}) \div \sup_{A\in\mathcal{A}, B\in\mathcal{B}} D_{\rho_\phi}(A, B); \quad \nu_{D_{\rho_\phi}}(\mathcal{A}, \mathcal{A}) \div \sup_{A\in\mathcal{A}} D_{\rho_\phi}(A, A) = 0,$$

is an ε-approach merotopy on X. In terms of the labelled sets $\mathfrak{I}_1, \mathfrak{I}_2, \mathfrak{I}_3, \mathfrak{I}_4, \mathfrak{I}_5, \mathfrak{I}_6$ in Fig. 2, we can identify the following lower associated set E_ε in (Assoc.1) and upper associated E^ε in (Assoc.2) with respect to $\nu_{D_{\rho_\phi}}$.
(**Assoc**.1) $E_\varepsilon(\mathfrak{I}_1) = \{\mathfrak{I}_2, \mathfrak{I}_3, \mathfrak{I}_4, \mathfrak{I}_5, \mathfrak{I}_6\}$, where

$$\nu_{D_{\rho_\phi}}(\mathfrak{I}_1, \mathfrak{I}_i) < \varepsilon \text{ for } i \in \{2, 3, 5, 6\},$$

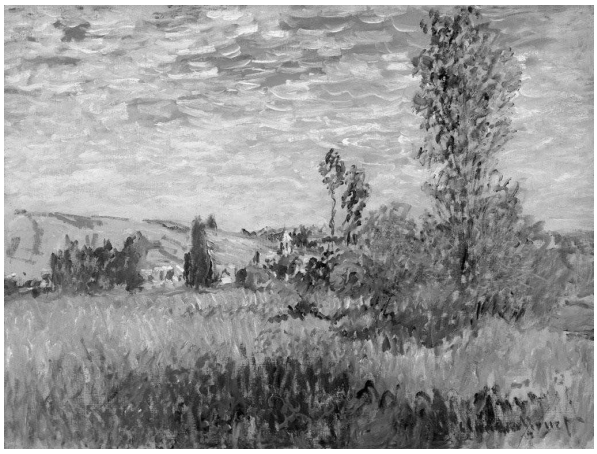
i.e., for $a \in \mathfrak{I}_1, b \in \mathfrak{I}_i, i \neq 1$, $\rho_\phi(a, b) < \varepsilon$, since the colours of all of the pixels are similar in each set $\mathfrak{I}_i \in E_\varepsilon(\mathfrak{I}_1)$ in Fig. 2. The sets in $E_\varepsilon(\mathfrak{I}_1)$ are sufficiently near $\mathfrak{I}_1$.
(**Assoc**.2) $E^\varepsilon(\mathfrak{I}_4) = \{\mathfrak{I}_7\}$, where

$$\nu_{D_{\rho_\phi}}(\mathfrak{I}_4, \mathfrak{I}_7) > \varepsilon,$$

i.e., for $a \in \mathfrak{I}_4, b \in \mathfrak{I}_7$, $\rho_\phi(a, b) > \varepsilon$, due to the fact that the green colour of each the pixels in $\mathfrak{I}_4$ is dissimilar to the purple or white colour of the pixels in $\mathfrak{I}_7$ in Fig. 2. In effect, the sets in $E^\varepsilon(\mathfrak{I}_4)$ are far apart from $\mathfrak{I}_4$ with respect to $\nu_{D_{\rho_\phi}}$.

Example 5.3. Sufficiently Near Strong p-Containment
After a manner similar to Example 4.3, let (X, ν_{ρ_ϕ}) be an ε-approach nearness on X and $p \div$ 'sufficiently near' defined for $\{A\}, \{B\} \in \mathcal{P}(X)$ such that $\nu_{\rho_\phi}(\{A\}, \{B\}) < \varepsilon$. Consider $\mathfrak{I}_1, \mathfrak{I}_2, \mathfrak{I}_3$ in Fig. 2. It is a straightforward task to verify that
(**p**.1) $\mathfrak{I}_1 \underset{p}{\subset} \mathfrak{I}_2 \subset \mathfrak{I}_3$ implies $\mathfrak{I}_1 \underset{p}{\subset} \mathfrak{I}_3$,
(**p**.2) $\mathfrak{I}_1 \subset \mathfrak{I}_2 \underset{p}{\subset} \mathfrak{I}_3$ implies $\mathfrak{I}_1 \underset{p}{\subset} \mathfrak{I}_3$,
(**p**.3) Considering only $\mathfrak{I}_1, \mathfrak{I}_2, \mathfrak{I}_3$,

$$\mathfrak{I}_1 \underset{p}{\subset} \mathfrak{I}_3 \text{ and} \mathfrak{I}_2 \underset{p}{\subset} \mathfrak{I}_3 \text{ implies } \bigcup_{i=1}^{2} \mathfrak{I}_i \underset{p}{\subset} \mathfrak{I}_3.$$



3.1: Monet meadow

3.2: Nbd $N_{r_{\phi_{grey}}}$

Figure 3: Sample Monet Meadow nbd $N_{r_{\phi_{grey}}}$, with $r_{\phi_{grey}}$=10

Example 5.4. Visual Neighbourhood in a Digital Image
Consider visual neighbourhoods in digital images, where each point is an image pixel (picture element). A pixel is described in terms of its feature values. Pixel features include grey level intensity and primary colours red, green, and blue with wavelengths 700 nm, 546.1 nm and 435.8 nm, respectively)[5], texture, and shape information. Visual information (feature values) is extracted from each pixel with a variety of probe functions.

For example, consider a xixth century, St. Martin, Vetheuil landscape by C. Monet rendered as a greyscale image in Fig. 3.1. Let $\phi_{grey}(x)$ denote a probe that extracts the greylevel intensity from a pixel x and let $r_{\phi_{grey}} = 10$. This will lead to the single visual neighbourhood represented by the green-shaded regions shown in Fig. 3.2. To obtain the visual nbd in Fig. 3.2, replace the greylevel intensity of each point sufficiently near the intensity $\phi_{grey}(x_0)$ with a green colour. The result is green-coloured visual nbd $N_{r_{\phi_{grey}}}$ in Fig. 3.2. This set of intensities in the visual nbd shown in $N_{r_{\phi_{grey}}}$ is an example of an open set contain numbers representing intensities that are sufficiently near x_0. To verify this, notice that the pixel intensities for large regions of the sky, hills and meadow in Fig. 3.1 are quite similar. This is the case with the sample pixels (points of light) x_0, x_1, x_2 in Fig. 3.2, where the in $\left|\phi_{grey}(x_0) - \phi_{grey}(x_1)\right| < r_{\phi_{grey}}$ and $\left|\phi_{grey}(x_0) - \phi_{grey}(x_2)\right| < r_{\phi_{grey}}$.

In summary, the lower associated set $E_\varepsilon(\{\{x_i\}\})$ is the set of all visual neighbourhoods of the pixel x_i in Fig. 3.2 that are descriptively ε-near each other. In addition, one can also observe

[5]The amounts of red, green and blue that form a particular colour are called *tristimulus* values. Let R, G, B denote red, green, blue tristimulus values, respectively, with green almost in the middle of the wavelengths of the visual spectrum, which is at 568 nm. Then define the following probe functions to extract the colour components of a pixel.

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = 1 - r - g.$$

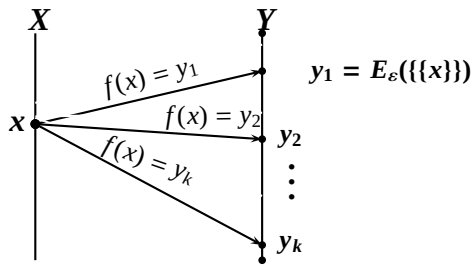


Figure 4: $f(x) = |E_\varepsilon(\{\{x\}\})| > 0$

that the upper associated set $E_\varepsilon(\{\{x_i\}\})$ contains all visual neighbourhoods that are descriptively dissimilar to x_i.

Example 5.5. Bipartite Graph for Associated Sets
Although this example continues the discussion of paintings, the proposed bipartite graph representation of associated sets is easily extended to members of any pair of nonempty sets. For example, consider classifying paintings by a particular artist by collecting together nonempty associated lower sets of sufficiently near neighbourhoods extracted from pairs of pictures. To see this, let X denote a set of query images and let Y denote a set of test images (*i.e.*, X contains pictures showing paintings, where each painting in X is compared with the paintings in the set of sample paintings Y).

The goal is to collect together those pictures in Y containing neighbourhoods of points in $y \in Y$ that are sufficiently similar to neighbourhoods of points in a picture $x \in X$. Let $N_a \in X, N_b \in Y$ denote neighbourhoods that are sufficiently near. Then construct the lower associated set $E_\varepsilon(N_a) = \{N_b, \dots\}$. A query image is similar to a test image if, and only if, $E_\varepsilon(N_a) > 0$.

Given approach spaces $(X, \nu_{D_{\rho_{\phi_{grey}}}}), (Y, \nu_{D_{\rho_{\phi_{grey}}}})$, consider a function $f : X \to Y$ defined by $f(x) = |E_\varepsilon(x)|$, where $x \in X$. Then the relation between a particular painting and one or more associated lower sets can be represented by a bipartite graph (see Fig. 4). The image set

$$O = \{f(x_i) : i \in \text{ and } |f(x_i)| > 0\}$$

can be extracted from Fig. 4. The set O has interest, since two of its members reveal the least similar and most similar paintings in relation to a particular query image. That is, $\inf\{O\}, \sup\{O\}$ function values correspond to the least similar and most similar of the paintings that are sufficiently near the query image $x \in X$.

Similarly, one can determine the collection of those paintings dissimilar to a given query picture with a nonempty associated upper set containing visual neighbourhoods taken from the query image and a test image.

References

Agronsky, S.J. (1982). A generalization of a theorem of maximoff and applications. *Trans. Amer. Math. Soc.* **273**(2), 767–779.

Beer, G., A. Lechnicki, S. Levi and S.A. Naimpally (1992). Distance functionals and suprema of hyperspace topologies. *Annali di Matematica pura ed applicata* **CLXII**(IV), 367–381.

Bourbaki, N. (1966). *Elements of Mathematics. General Topology, Part 1*. Hermann & Addison-Wesley. Paris & Reading, MA, U.S.A. i-vii, 437 pp.

Bourbaki, N. (1971). *Topologie générale, 1-4*. Hermann. Paris. Springer-Verlag published a new edition in 2007.

Bruckner, A.M. (1967). On characterizing classes of functions in terms of associated sets. *Canad. Math. Bull.* **10**(2), 227–231.

Čech, E. (1966). *Topological Spaces, revised Ed. by Z. Frolik and M. Katătov*. John Wiley & Sons. NY.

Coble, A.B. (1922). Associated sets of points. *Trans. Amer. Math. Soc.* **24**(1), 1–20.

Engelking, R. (1989). *General Topology, Revised & completed edition*. Heldermann Verlag. Berlin.

Hausdorff, F. (1914*a*). *Grundzüge der Mengenlehre*. Veit and Company. Leipzig. viii + 476 pp.

Hausdorff, F. (1914*b*). *Set Theory*. AMS Chelsea Publishing. Providence, RI. 352 pp.

Henry, C.J. (2010). Near Sets: Theory and Applications, Ph.D. dissertation, supervisor: J.F. Peters. PhD thesis. Department of Electrical & Computer Engineering.

Hocking, J.G. and G.S. Young (1988). *Topology*. Dover. NY.

Khare, M. and S. Tiwari (2010). Grill determined *L*-approach merotopological spaces. *Fund. Inform.* **48**, 1–12.

Khare, M. and S. Tiwari (2012). *L*-approach merotopies and their categorical perspective. *Demonstratio Math.* **45**(3), 699–716.

Leader, S. (1959). On clusters in proximity spaces. *Fundamenta Mathematicae* **47**, 205–213.

Lowen, R. (1997). *Approach Spaces: The Missing Link in the Topology-Uniformity-Metric Triad*. Oxford Mathematical Monographs, Oxford University Press. Oxford, UK. viii + 253pp.

Lowen, R., D. Vaughan and M. Sioen (2003). Completing quasi metric spaces: an alternative approach. *Houstan J. Math.* **29**(1), 113–136.

Naimpally, S.A. and J.F. Peters (2013). *Topology with Applications. Topological Spaces Via Near and Far*. World Scientific. Singapore.

Peters, J. F. and S. Tiwari (2012). Completing extended metric spaces: an alternative approach. *Appl. Math. Letters* **25**, 1544–1547.

Peters, J.F. (2011). How near are Zdzisław Pawlak's paintings? Merotopic distance between regions of interest. pp. 1–19. , *in* A. Skowron, S. Suraj, Eds., Intelligent Systems Reference Library volume dedicated to Prof. Zdzisław Pawlak, Springer, Berlin.

Peters, J.F. (2013). Local near sets: Pattern discovery in proximity spaces. *Math. Comp. Sci.* **7**(1), *to appear*.

Peters, J.F. and S. Tiwari (2011). Approach merotopies and near filters. *Gen. Math. Notes* **3**(1), 32–45.

Peters, J.F. and S.A. Naimpally (2012). Applications of near sets. *Amer. Math. Soc. Notices* **59**(4), 536–542.

Petrakiev, I. (2009). On self-associated sets of points in small projective spaces. *Comm. in Algebra* **37**(2), 397–405.

Tiwari, S. (Jan. 2010). Some Aspects of General Topology and Applications. Approach Merotopic Structures and Applications, supervisor: M. Khare. PhD thesis. Department of Mathematics, Allahabad (U.P.), India.

Zahorski, Z. (1950). Sur la première dérivée. *Trans. Amer. Math. Soc.* **69**, 1–54.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 13–31

Unsupervised Detection of Outlier Images Using Multi-Order Image Transforms

Lior Shamir[a,*]

[a]*Lawrence Technological University, 21000 W Ten Mile Rd., Southfield, MI 48075, United States.*

Abstract

The task of unsupervised detection of peculiar images has immediate applications to numerous scientific disciplines such as astronomy and biology. Here we describe a simple non-parametric method that uses multi-order image transforms for the purpose of automatic unsupervised detection of peculiar images in image datasets. The method is based on computing a large set of image features from the raw pixels and the first and second order of several combinations of image transforms. Then, the features are assigned weights based on their variance, and the peculiarity of each image is determined by its weighted Euclidean distance from the centroid such that the weights are computed from the variance. Experimental results show that features extracted from multi-order image transforms can be used to automatically detect peculiar images in an unsupervised fashion in different image datasets, including faces, paintings, microscopy images, and more, and can be used to find uncommon or peculiar images in large datasets in cases where the target image of interest is not known. The performance of the method is superior to general methods such as one-class SVM. Source code and data used in this paper are publicly available, and can be used as a benchmark to develop and compare the performance of algorithms for unsupervised detection of peculiar images.

Keywords: Outlier detection, peculiar images, image analysis, image transform, multi-order transforms.
2010 MSC: 68T10, 62H35, 68T45, 62H30 .

1. Introduction

Unsupervised detection of peculiar images is the ability of a computer system to automatically detect images that are different from the other "regular" images in an image dataset. While in tasks such as image classification the system can be trained in a supervised fashion using "ground truth" samples, in unsupervised detection of peculiar images the algorithm cannot rely on data samples or models that reflect the "regular" images or the target images of interest.

The problem of detecting data points significantly different from the other data is often referred to as *outlier detection* (Hodge & Austin, 2004). Many established algorithms consider outlier

*Corresponding author
Email address: lshamir@mtu.edu (Lior Shamir)

detection as a by-product of clustering algorithms by searching for background noise samples that do not belong in a cluster (Aggarwal & Yu, 2000; Guha, Rastogi & Shim, 2001). Other methods are based on searching for samples that do not belong in a cluster and are also not background noise, but are substantially different from the other samples in the dataset (Breunig et al., 2000; Knorr & Ng, 1999; Fan et al., 2006). While many of the outlier detection algorithms were designed and tested using lower dimensionality, other methods aim at automatic outlier detection in higher dimensionality data (Aggarwal & Yu, 2001; Roth, 2005; Fan, Cehn & Lin, 2005; Lukashevich, Nowak & Dunker, 2009). Applications of outlier detection include credit card fraud, network intrusion detection, surveillance, financial applications, cell phone fraud, safety critical systems, loan application processing, defect detection in factory production lines, and sensor networks (Zhang et al., 2007).

While outlier detection has been studied in the context of a broad range of applications, less work has yet been done on unsupervised detection of peculiar images in image datasets. Here we describe a generic method that can be used for automatic detection of peculiar images in image datasets based on a large set of image content descriptors extracted from the raw pixels, image transforms, and compound image transforms. Applications include, for instance, the search for peculiar cells or tissues in large datasets of microscope images, which can be used to detect phenotypes of particular scientific interest (d'Onofrio & Mango, 1984; Carpenter, 2007; Jonesa et al., 2009).

When the target image is known, the task of detecting a peculiar image can be related to the problem of Content-Based Image Retrieval, and numerous effective methods of measuring similarities between images in the context of CBIR have been proposed (Bilenko, Basu, & Mooney, 2004; Kameyama et al., 2006). However, since in this study the detection of a peculiar image in an image dataset should be done automatically in an unsupervised manner, no assumptions can be made neither about the target image nor about the context of the images in the data base. That is, the computer system should automatically characterize the "typical" image in the dataset, and detect images that are different from it. Since no pre-defined model of the data can be used, effective systems for unsupervised automatic detection of peculiar images need to extract different image features that will cover different aspects of the image content, and thus be able to characterize and analyze a broad spectrum of image data.

Here we use a large set of image content descriptors extracted from the raw images, image transforms, and multi-order transforms, and apply a statistical analysis to weight the different image features by their ability to reflect the data and detect peculiar images. The primary advantage of the method is its generality, which makes it effective for the analysis of a broad variety of image datasets without the need for tuning or adjustments. In Section 2 we briefly describe the set of image features and multi-order transform model used in this study, in Section 3 we describe the unsupervised detection of the peculiar images, in Section 4 the performance evaluation method of the proposed algorithm is discussed, and in Section 5 the experimental results are presented.

2. Image features

The set of image content descriptors used in this study is based on the feature set used by the *wndchrm* algorithm, which is a large set of numerical image content descriptors that cover a broad

range of aspects of the visual content (Shamir et al., 2008a; Orlov et al., 2008; Shamir et al., 2010). Basically, the *wndchrm* feature set includes several generic image features such as high-contrast features (object statistics, edge statistics, Gabor filters), textures (Haralick, Tamura), statistical distribution of the pixel values (multi-scale histograms, first four moments), factors from polynomial decomposition of the image (Chebyshev statistics, Chebyshev-Fourier statistics, Zernike polynomials), Radon features, and fractal features. A detailed description of these image content descriptors and the way they are used in the context of the *wndchrm* feature set is available in (Shamir et al., 2008a; Orlov et al., 2008; Shamir et al., 2010; Shamir, 2008; Shamir et al., 2009). The reason for using a large set of features is that the search for peculiar images is unsupervised, and no assumptions can be made regarding the possible difference between the peculiar and non-peculiar images. Therefore, it is important that the set of image content descriptors is comprehensive enough so that at least some of the image features will be likely to sense differences between a regular and a peculiar image in a given image dataset.

As will be discussed in Section 5, a key contributor to the ability of the method proposed in this paper to detect peculiar images in an unsupervised fashion is the extraction of the image content descriptors not just from the raw pixels, but also from image transforms and compound image transforms. The extraction of image features from compound image transforms has been shown to contribute significantly to the performance of general-purpose image classifiers (Shamir et al., 2008a; Orlov et al., 2008; Shamir et al., 2010, 2009), and can therefore be effective for peculiar image detection in cases where the differences between the typical and the peculiar image should be determined automatically, without using "ground truth" samples or any other prior knowledge about the data. The image transforms include the Fourier, Chebyshev, Wavelet, and the edge-magnitude transform, as well as multi-order transform combinations. The combinations of transforms include the Fourier transform of the Chebyshev transform, the wavelet transform of the Chebyshev transform, the Fourier transform of the wavelet transform, the wavelet transform of the Fourier transform, the Chebyshev transform of the Fourier transform, and the Fourier and Chebyshev transforms of the edge magnitude transform. A detailed description of the tandem transform combinations can be found in (Shamir et al., 2010; Shamir, 2008), and the total number of images features extracted using these transforms is 2659 (Shamir et al., 2008a; Shamir, 2008). The length of the chain of transforms is limited to the first and second order of the image transforms, as experiments showed that using compound transforms with order higher than two typically does not contribute to the informativeness of the image analysis system (Shamir et al., 2009). The effect of using the multi-order image transforms on the ability of the algorithm to automatically detect peculiar images will be discussed in Section 5.

For color images we used a color transform, which is based on transforming the RGB pixels into the HSV space, followed by classification of the HSV triplets into one of 16 color classes using fuzzy logic modeling of the human perception of these colors (Shamir et al., 2006). Then, the Fourier, Chebyshev, and wavelet transforms of the color transform are computed, and the set of image features is extracted as described in (Shamir et al., 2010). When the color transform is also used, the total number of image features is 3658 (Shamir et al., 2010). Figure 1 illustrates the paths of the transforms and compound transforms used by the feature set.

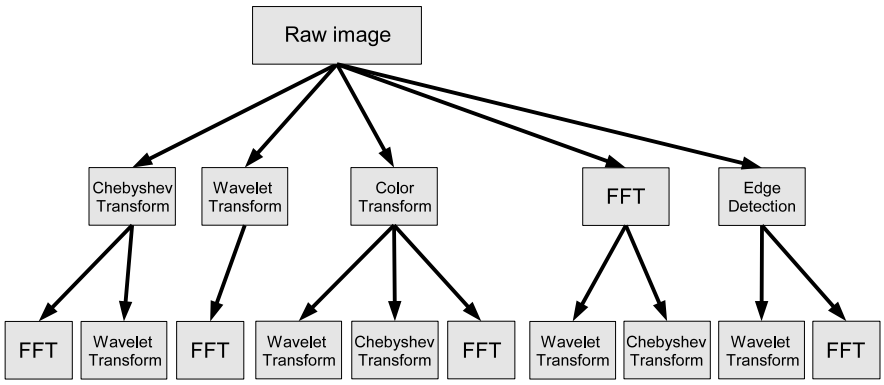


Figure 1. Paths of multi-order image transforms.

3. Automatic detection of peculiar images

In order to automatically detect peculiar images, it is first required to characterize the "typical" image in the dataset. Since many image features are used without prior knowledge about the dataset, it can be assumed that not all content descriptors are relevant to the image dataset at hand, and might represent noise. Therefore, it is required to select the image features that are the most informative, and can potentially discriminate between peculiar and non-peculiar images.

In the first stage of the algorithm, all image features are normalized to the interval [0, 1], so that the differences between the values of different image features can be compared without introducing a numerical bias. For instance, if the values of one image feature are in the range of [0, 1000] while the values of another are in the range of [0, 10], a numerical difference of 5 between the values of the first feature extracted from two different images can be considered small, while the same numerical difference can be much more substantial for the second feature, in which it is half of the entire range.

In the next step, the mean, median, and variance of each image feature are computed. To characterize the "typical" feature values of an image in the dataset, the highest 5% and the lowest 5% of the values of each image feature are ignored when computing the mean and standard deviation, so that extreme values that results from noise, artifacts, or peculiar images will not affect the mean and variance of the "typical" images.

After these values are computed, each image in the dataset is compared to the "typical" image using Equation 3.1

$$D_i = \sum_{f \in F} (1 - \sigma_f)^k \cdot \frac{|f_i - \overline{f}|}{\sigma_f}, \tag{3.1}$$

where D_i is the dissimilarity value of image i from the "typical" image in the dataset, f is a feature in the feature set F, $\overline{f}$ is the median of the values of feature f in the given image dataset, f_i is the value of the feature f computed from the image i, σ_f is the standard deviation of feature f, and k is a constant value set to 25. The value of k will be thoroughly discussed in Section 5. The D_i dissimilarity value can be conceptualized as the sum of Z scores computed for each feature separately, such that each score's contribution to the total distance is inversely dependent on

the standard deviation. That is, features that have lower standard deviation are considered more "representative" features, while feature that their values are more sparsely distributed are assumed to provide a weaker representation of the "typical" image and are therefore assigned with a lower score and have a weaker affect on the dissimilarity value.

Clearly, image features that their values are constant across the dataset (and therefore $\sigma = 0$) cannot provide any useful information in this model, and can therefore be safely ignored without affecting the performance. On the other hand, the values of some of the other features can be sparsely distributed across the image dataset, and therefore the median of these values cannot be considered as a value that reliably represents the typical image. For that reason, the effect of each image feature is weighted by its standard deviation, which is used as an assessment of the feature's informativeness and its ability to characterize the typical image.

While in Equation 3.1 the effect of features that their values are sparsely distributed is weakened by using the standard deviation as a measurement of their informativeness, it can be assumed that many of the features will not be informative for a given image dataset at hand, and therefore the high number of irrelevant features can add noise to the analysis and negatively affect the performance. In order to reduce the effect of non-informative features, 90% of the features with the highest σ are ignored, and the remaining 10% are used by Equation 3.1 to compute the distance between a given image and the "typical" image in the dataset. Since the image features are also weighted by their standard deviation, the performance of this method is not highly sensitive to small changes in the number of features that are used, as will be discussed in Section 5. This approach of combining feature selection and feature weighting is conceptually similar to the approach of the feature selection in the *wndchrm* image classifier (Shamir et al., 2008a; Orlov et al., 2008; Shamir et al., 2010).

In many cases, using σ to assess the informativeness of the features and their ability to differentiate between a peculiar and a typical image might not be optimal and can lead to the sacrifice of some of the information. For instance, if the values of a certain image feature range between 0 and 0.8 for most images in the dataset, but is always 1 for a certain peculiar image, this feature could have been effectively used to detect the peculiar image, but will be assigned with a low score due to the sparse distribution of the values. On the other hand, features can be assigned with high scores due to the consistency of their values, while these image features might have little ability to differentiate between a typical and a peculiar image. Since the goal of the method described in this paper is to detect peculiar images in an unsupervised fashion, no assumptions or prior knowledge about the data can be used, and therefore the image content descriptors cannot be selected or scored based on a target peculiar image. However, by using a large set of image features, it can be expected that some of the features that are assigned with high scores will be able to differentiate between a peculiar and a typical image. This will be demonstrated in Section 5.

In summary, the following pseudo code summarizes the outlier detection algorithm:

Step 1: Compute image features for all images.

Step 2: Reject the lowest and highest 5% values of each feature.

Step 3: Compute the mean M_f of the values of each feature f.

Step 4: Compute the σ_f of the values of each feature f.

Step 5: Reject 90% of the features with the lowest σ.

Step 6: Compute d for each image I such that $d = \sqrt{\sum_f (1-\sigma_f)^k \cdot \frac{(I_f - M_f)}{\sigma_f}}$.

Step 7: Sort the images in the dataset by d.

The computational complexity of the algorithm is $O(F \cdot I \log I)$, where I is the number of images and F is the number of features computed for each image. The computational complexity is determined by the complexity of sorting all values of each feature, which is the most computationally demanding task in the algorithm described above. The bottleneck of the process, however, is the computational complexity of the Wndchrm feature set, which is much more complex as described in (Shamir et al., 2008a, 2009).

4. Performance evaluation

In order to test the performance of the proposed method, several different image datasets were used. These datasets include the Brodatz texture album (Brodatz, 1966), the COIL-20 object image collection (Nene, Nayar & Murase, 1996), JAFFE and AT&T face datasets (Samaria & Harter, 1994; Lynos et al., 1998), the MNIST handwritten digit collection (LeCun et al., 1998; Liu et al., 2003), and a dataset of digitized paintings of Van Gogh, Monet, Dali, and Pollock Shamir et al. (2010). Since the MNIST dataset contains a large number of images, a subset of 100 images from the first two classes (0 and 1) were used in the experiment. For microscopy images we used the CHO (Chinese Hamster Ovary) dataset (Boland & Murphy, 2001), consisting of fluorescence 512×382 microscopy images of different sub-cellular compartments, and the Pollen dataset (Duller et al., 1999), which is a dataset of 25×25 images of geometric features of pollen grains. The CHO dataset might not be considered a perfect representation of biological content (Shamir et al., 2011), but it is used in this study for general-purpose outlier detection. These two datasets are available for download as part of the IICBU-2008 benchmark suite at `http://ome.grc.nia.nih.gov/iicbu2008` (Shamir et al., 2008b), and sample images of the different classes are shown by Figures 2 and 3. The image datasets used in this study are listed in Table 1.

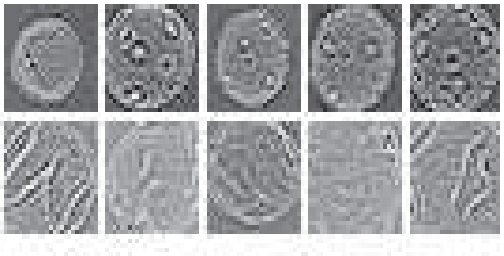


Figure 2. Sample images of class “obj_198” (top) and “obj_212” (bottom) taken from the pollen dataset.

Table 1. Image datasets used for the experiments.

dataset	typical class	peculiar class	images per class
Pollen	obj_198	obj_212	45
CHO	giantin	hoechst	69
JAFFE	KA	KL	22
AT&T	1	2	10
Painters 1	Pollock	Dali	30
Painters 2	Monet	Van Gogh	30
Brodatz 1	Bark	Brick	4
Brodatz 2	Wood	Wool	4
MNIST	0	1	100
COIL-20	obj1	obj2	71

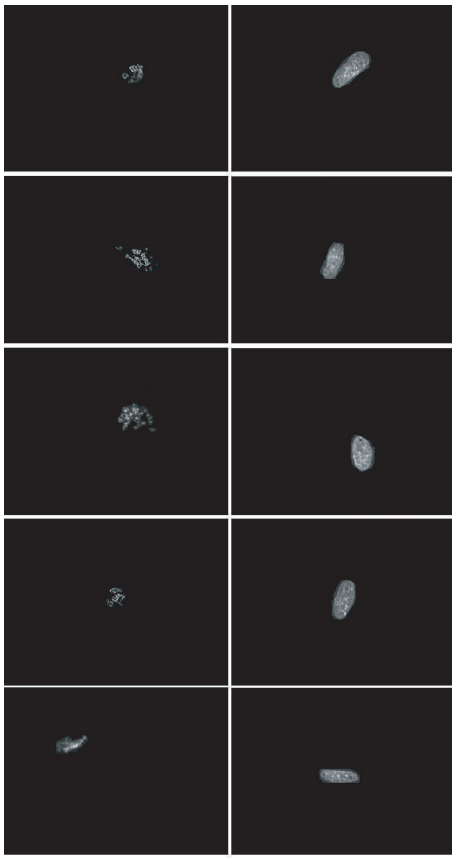


Figure 3. Sample images of giantin (left) and hoechst (right) taken from the CHO dataset.

As the table shows, these datasets were used such that two classes from each dataset were selected: one class was used as the "typical" class and the other as a pool of "peculiar" images.

In each run the tested dataset included all images from the typical class, and one image from the peculiar class. The experiment was repeated for each image in the peculiar class, such that in each run a different image from the peculiar class served as the peculiar image. For instance, the AT&T face dataset has 10 images in each class, and therefore it was tested 10 times such that in each run all 10 images of person *1* were used and one image of person *2* (a different image in each of the 10 runs). The goal of the algorithm was to automatically detect the single image of person *2* among the dataset of 11 images (10 images of person *1* and one image of person *2*).

The performance was evaluated by the number of times the algorithm correctly detected the peculiar image in the set (which included the "typical" images and the one "peculiar" image), divided by the total number of images in the peculiar class. Another performance metrics used in this study is the rank-10 detection accuracy, which was measured as the percentage of the cases in which the peculiar image was among the first 10 candidates with the highest dissimilarity value as determined by Equation 3.1.

5. Results

The performance of the automatic detection of peculiar images was evaluated as described in Section 4, and the rank-1 and rank-10 detection accuracies for each of the tested datasets are listed in Table 2.

Table 2. Rank-1 and rank-10 accuracy of the detection of the peculiar image.

Dataset	Rank-1 accuracy	Rank-10 accuracy
Pollen	29/45	34/45
CHO	57/69	69/69
Jaffe	16/22	22/22
AT&T	10/10	10/10
Painters 1	26/30	30/30
Painters 2	0/30	18/30
Brodatz 1	4/4	4/4
Brodatz 2	4/4	4/4
MNIST	29/100	92/100
COIL-20	38/71	71/71

As the table shows, in almost all cases the proposed algorithm was able to automatically detect the peculiar images in accuracy significantly better than random. For instance, with the Pollen dataset the algorithm was able to automatically find the peculiar image in 29 times out of 45 attempts (each attempt with a different image), and the rank-10 detection was accurate in 34 times. The noticeable exception is the second datasets of painters, which consists of paintings of Monet and Van Gogh. In that case, the proposed method was not able to automatically detect any of the tested Van Gogh paintings in a set of Monet paintings, and the rank-10 accuracy was 60%. Since Monet and Van Gogh were inspired from each other, their artistic styles are similar to each other, and it usually requires knowledge in art to differentiate between the works of the two painters.

The other painter dataset that was tested demonstrated a much higher detection accuracy since the two painters, Jackson Pollock and Salvador Dali, belong in different schools of art (Abstract Expressionism and Surrealism, respectively) and the differences between their styles are highly noticeable even without any previous knowledge or training in art.

The results specified in Table 2 demonstrate the generality of the method and its ability to handle very different image datasets in a fully automatic fashion, and without the need to select or tune parameters. The generality of the method can also be demonstrated by the different classes of the Pollen dataset. While the results in Table 2 are based on obj_198 as the "typical" class and "obj_212" as the peculiar class, the pollen dataset includes seven classes (Shamir et al., 2008b). Table 3 shows the rank-1 detection accuracy of all combinations of the seven classes in the pollen dataset, such that each cell is the detection accuracy when the row the "typical" class and the column is the "peculiar" class. As the table shows, the detection accuracy is significantly higher than random in all combinations of "typical" and "peculiar" classes, demonstrating the generality of the proposed method.

Table 3. Rank-1 detection accuracy (%) of all combinations of typical and peculiar classes using the pollen dataset.

Regular\Peculiar	198	212	216	360	361	405	406
198	-	64	51	67	53	69	67
212	55	-	55	65	58	67	67
216	63	67	-	63	58	64	64
360	61	64	67	-	64	72	69
361	57	67	65	67	-	72	69
405	66	71	73	69	67	-	71
406	63	59	65	69	69	61	-

As discussed in Section 3, 90% of the image features with the highest σ are ignored. Changing the number of features that are rejected and not used by the image dissimilarity evaluation of Equation 3.1 can change the dissimilarity value determined by the Equation for each image, and consequently affect the performance of the algorithm. Figures 4 and 5 show the rank-1 and rank-10 detection accuracy of the peculiar images when the number of used features is changed.

As the graphs show, while the peculiar image detection accuracy of some image datasets peaks when 10% of the features are used, in other datasets such as MNIST or COIL-20 the detection accuracy peaks when 40% of the features are used. In the case of MNIST, the rank-1 detection accuracy was elevated from 29% when 10% of the features were used to 94% with 40% of the image features. This shows that the detection of peculiar images can be optimized if the number of used image features is adjusted for the specific dataset. However, since the detection of the peculiar image is unsupervised, and in many cases the target peculiar image is unknown, adjusting the parameters for optimizing the performance based on sample target peculiar images might not be possible, and it is therefore required to use a general pre-defined parameter setting as was done for the performance figures reported in Table 2.

Another value that was determined experimentally is the K values in Equation 3.1. Figures 6

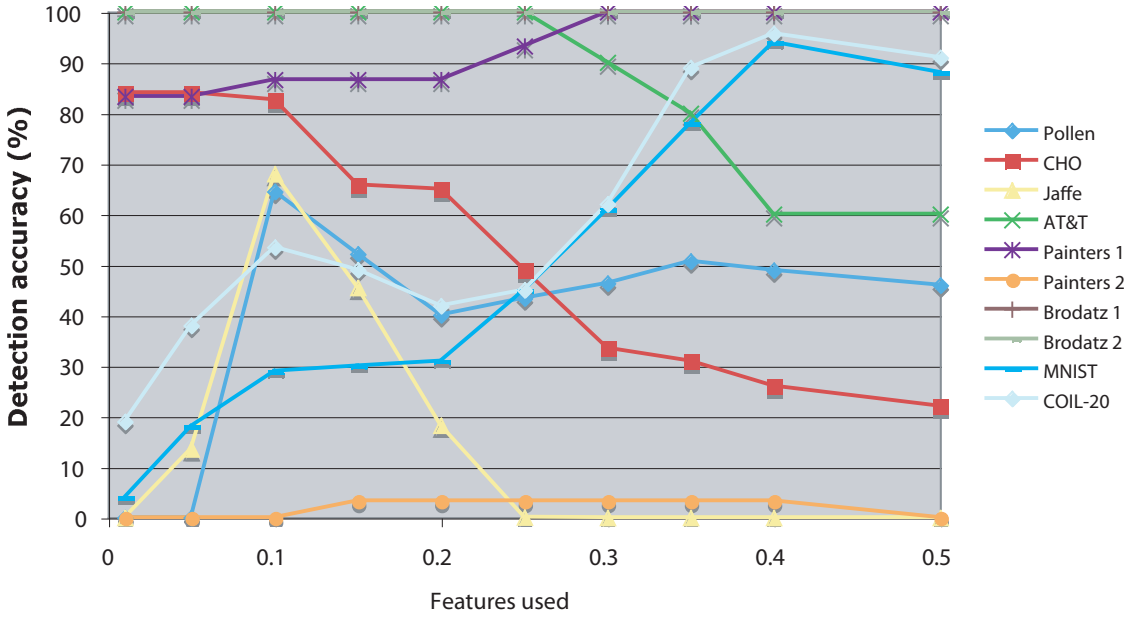


Figure 4. The rank-1 detection accuracy of the peculiar image as a function of the amount of features used.

and 7 show the rank-1 and rank-10 detection accuracy of the peculiar image as a function of the value of this parameter. In all cases, 10% of the image content descriptors were used as described in Section 3.

As the graphs show, the detection accuracy of the CHO dataset drops as the value of K increases, and the detection of a Dali painting in a set of Jackson Pollock paintings also peaks when the value of K is low. However, in most cases the detection accuracy of the peculiar image reaches its maximum when the value of K is around 25. In some of the tested image datasets, such as the Brodatz texture datasets and the rank-10 of the *Painters 1* and COIL-20 datasets, the detection accuracy of the peculiar image remained perfect regardless of the value of K.

A single peculiar image is expected to be detected more easily among a smaller dataset of regular images. That is, finding a peculiar image hidden in a dataset of millions of images is expected to be a more difficult task than finding a peculiar image in a dataset of just a dozen regular images. On the other hand, the presence of a large number of regular images allows the algorithm to find the image features that can discriminate between peculiar and regular images, and better estimate the weights of the features by their informativeness and ability to discriminate between a regular and a peculiar image as described in Section 3. To study the effect of the number of the regular images in the dataset we used the MNIST dataset, which provides a sufficient number of images of handwritten digits "0" and "1". Figure 8 shows the rank-1 and rank-10 detection accuracy of a peculiar image when changing the number of regular images in the dataset using 1000 images of the handwritten digit "0".

As the graph shows, the detection accuracy generally drops as more regular images are added

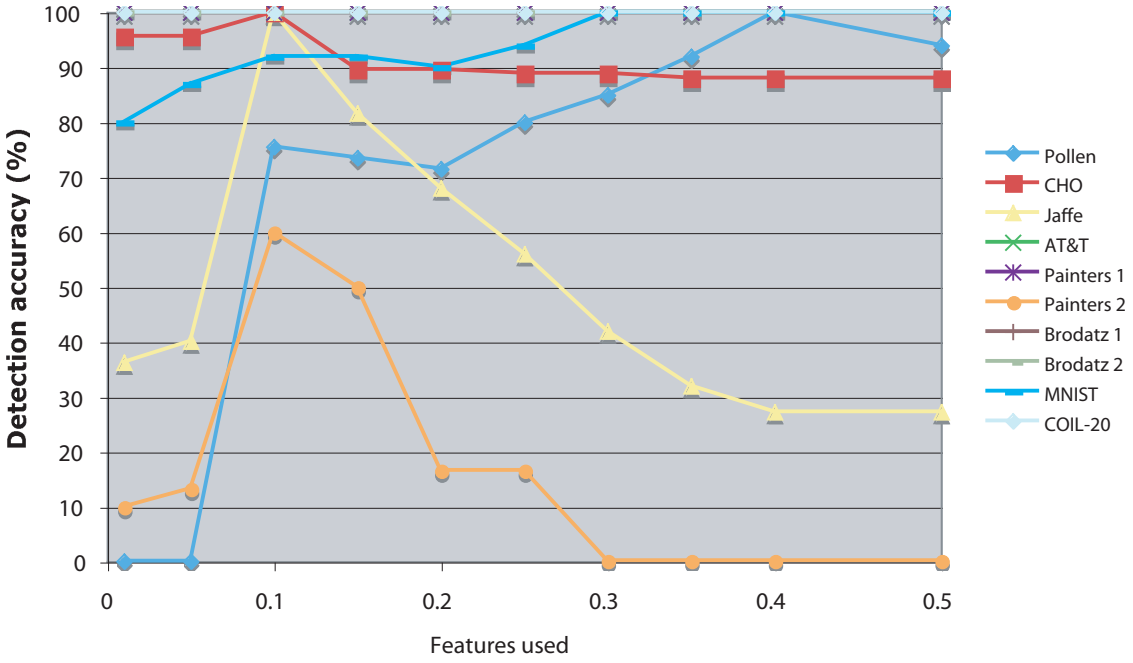


Figure 5. The rank-10 detection accuracy of the peculiar image as a function of the amount of features used.

to the dataset. Clearly, this is due to the lower difficulty of finding a peculiar image in a dataset of 50 images compared to correctly finding a single peculiar image among a dataset of 1000 regular images. However, while the rank-10 accuracy decreases as the number of regular images gets higher, the rank-1 detection accuracy drops to ~30% at around 80 regular images, but marginally changes when more regular images are added to the dataset. This can be due to the effect of the better weights assigned to the image features when the number of non-peculiar images in the dataset increases, which improve the ability of the algorithm to characterize the "typical" image in the dataset and differentiate it from peculiar images.

It should be noted, however, that while a higher number of regular images improves the feature weights, it also increases the probability that one the regular images in the dataset will be assigned with a high dissimilarity value computed by Equation 3.1. Since the algorithm aims to detect the irregular images in an unsupervised fashion, any difference between one of the images in the dataset and the "typical" image might lead to the detection of that image as "peculiar". For instance, in the MNIST dataset of the handwritten digit "0" the 10 most common images that were detected by the proposed algorithm as peculiar are shown in Figure 9.

As the figure shows, some of these handwritten digits are noticeably different from a standard handwritten digit "0". For instance, the top left digit has a black dot near it, while other images of handwritten "0" feature incomplete circles or thick lines. These images can confuse the algorithm since they are different from the typical image of the digit "0". Repeating the same experiment with a dataset of 100 manually selected "0" images that seemed relatively uniform led to perfect

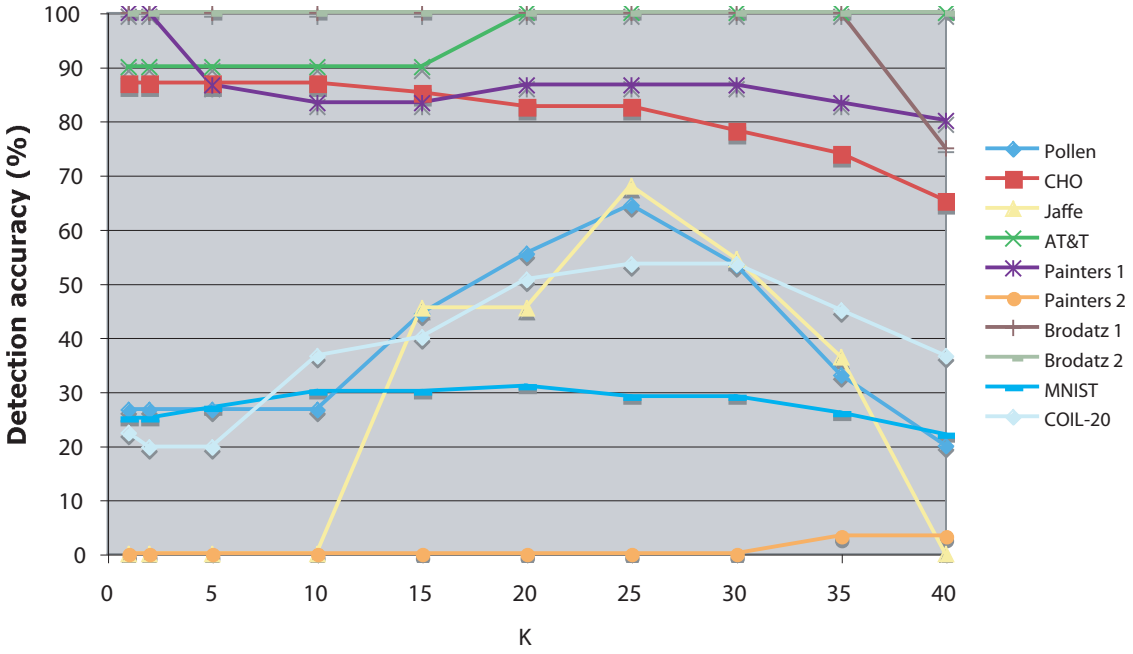


Figure 6. The rank-1 detection accuracy of the peculiar image as a function of K.

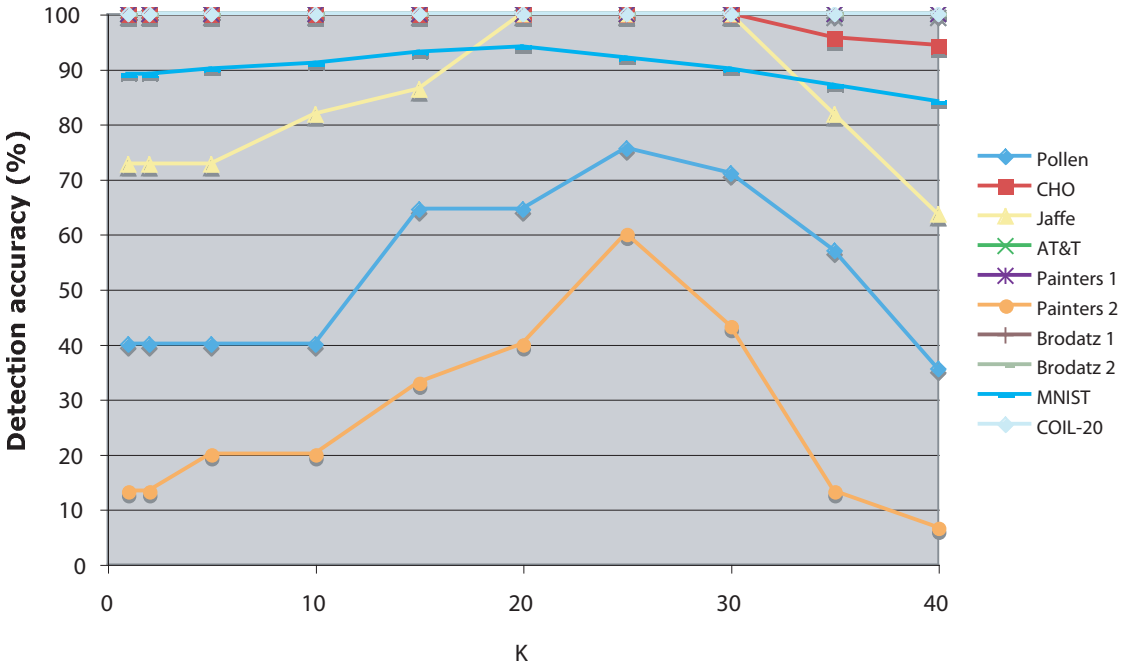


Figure 7. The rank-10 detection accuracy of the peculiar image as a function of K.

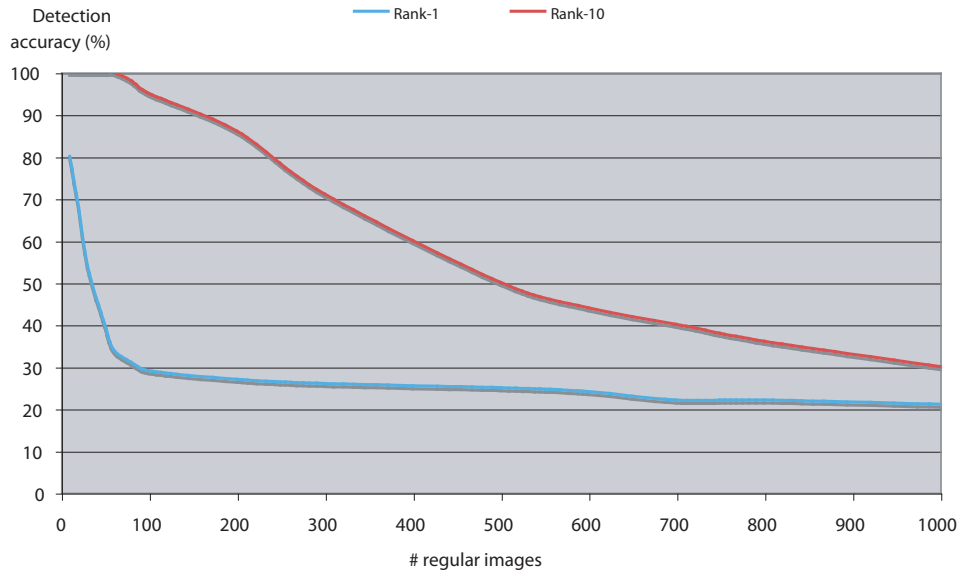


Figure 8. Rank-1 and rank-10 accuracy as a function of the number of regular images using the MNIST handwritten digits dataset using 10% of the image features.

Figure 9. The 10 most different images in the dataset of 1000 regular images of handwritten "0" digit.

detection accuracy of the "1" images.

Similarly, the most peculiar images of the CHO dataset and the AT&T dataset are shown in Figures 10 and 11, respectively. As the figures show, in the AT&T dataset the images are relatively similar to each other, and it is difficult to identify specific images that are significantly more different from the rest of the images. However, in the CHO dataset the peculiar images are noticeably different from the "typical" giantin images showed in Figure 3.

As showed by Figure 5, the detection of an image of the handwritten digit "1" in a large set of images of the handwritten digit "0" can be improved when using 40% of the image features. Figure 12 shows the detection accuracy when 40% of the image content descriptors are used.

As the figure shows, when using 40% of the features the detection accuracy also drops as the number of regular images gets larger, but the detection accuracy is significantly higher compared to the detection accuracy when using just 10% of the image content descriptors. The rank-10

Figure 10. The 10 most different images in the AT&T dataset. The leftmost image is the most peculiar.

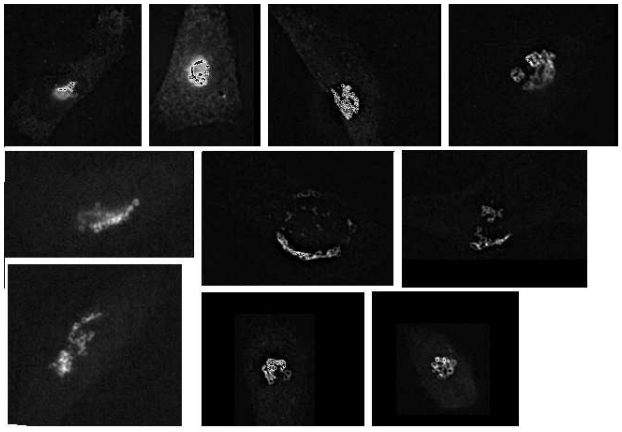


Figure 11. The 10 most different images in the CHO (giantin) dataset. The upper left image is the most peculiar and the lower right image is the least peculiar of the 10 samples.

accuracy, however, remains steady at 100% regardless of the number of regular images among which the peculiar image should be detected.

A key element in the proposed algorithm is the use of the large set of image features extracted from the raw pixels, but also from image transforms and compound transforms. To test the contribution of the features extracted from transforms and compound transforms, the performance of the proposed method was tested using features extracted from the raw pixels only, raw pixels and transforms, and raw pixel, transforms, and transforms of transforms. Table 4 shows the rank-1 and rank-10 detection accuracy when using image features computed using the raw pixels alone, and Table 5 shows the performance of the method when using also the image features extracted from the first-order image transforms. Table 2 shows the detection performance when using the raw pixels, image transforms, and transforms of transforms.

As the tables show, the use of image features extracted from transforms and multi-order image transforms has a significant effect on the performance of the method, and demonstrates the informativeness of standard image features extracted not just from the raw pixels, but also from image transforms and compound transforms (Rodenacker & Bengtsson, 2003; Gurevich & Koryabkina, 2006; Shamir et al., 2010, 2009).

5.1. Comparison the previous methods

The performance of the peculiar image detection method was also compared to the performance of one-class SVM (Scholkopf et al., 2001). The experiments were done with the LibSVM

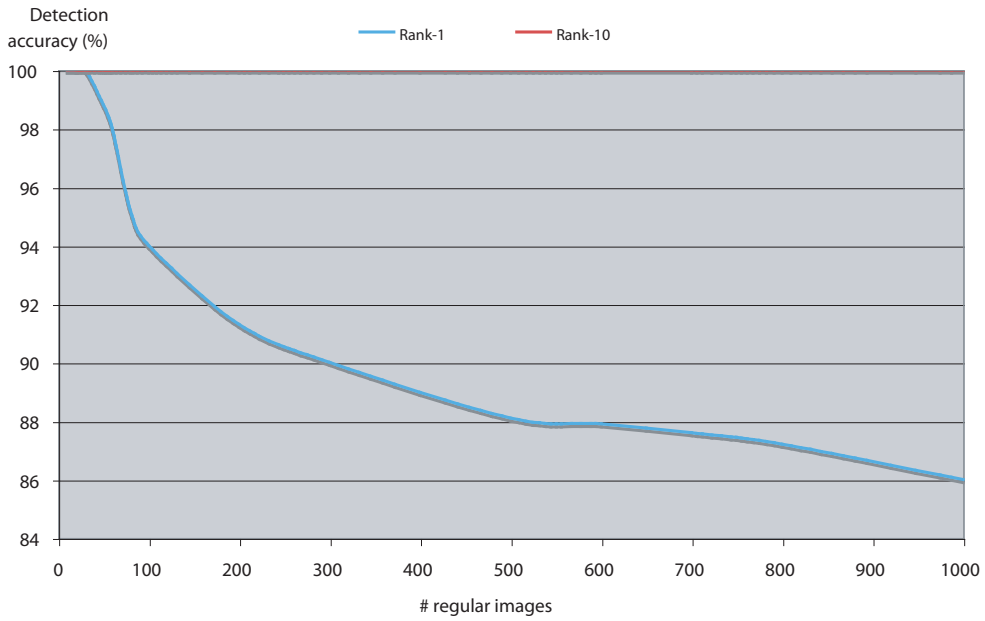


Figure 12. Rank-1 and rank-10 accuracy as a function of the number of regular images using the MNIST handwritten digits image dataset using 40% of the image features.

Table 4. Rank-1 and rank-10 accuracy of the peculiar image detection when using image features extracted from the raw pixels only.

Dataset	Rank-1 accuracy	Rank-10 accuracy
Pollen	11/45	19/45
CHO	21/69	36/69
Jaffe	9/22	14/22
AT&T	4/10	10/10
Painters 1	16/30	23/30
Painters 2	0/30	14/30
Brodatz 1	4/4	4/4
Brodatz 2	4/4	4/4
MNIST	9/100	56/100
COIL-20	11/71	44/71

support vector machine library using the "one-class" option with RBF (γ=5) and polynomial (d=5) kernels, where *nu* was set to 0.5 (Scholkopf et al., 2001; Fan, Cehn & Lin, 2005). The value K in Equation 3.1 was set to 25. Figure 13 shows the rank-1 detection accuracies using the method described in this paper and the one-class SVM with the two kernels.

As the graph shows, the detection using weighted distances from the means as described in this paper is substantially better compared to one-class SVM. The better performance when using the weighted distances from the means can be explained by the ability of the weighted feature space to

Table 5. Rank-1 and rank-10 accuracy of the peculiar image detection when using image features extracted from the raw pixels and image transforms.

Dataset	Rank-1 accuracy	Rank-10 accuracy
Pollen	18/45	25/45
CHO	39/69	53/69
Jaffe	13/22	22/22
AT&T	8/10	10/10
Painters 1	22/30	26/30
Painters 2	0/30	17/30
Brodatz 1	4/4	4/4
Brodatz 2	4/4	4/4
MNIST	29/100	92/100
COIL-20	38/71	68/71

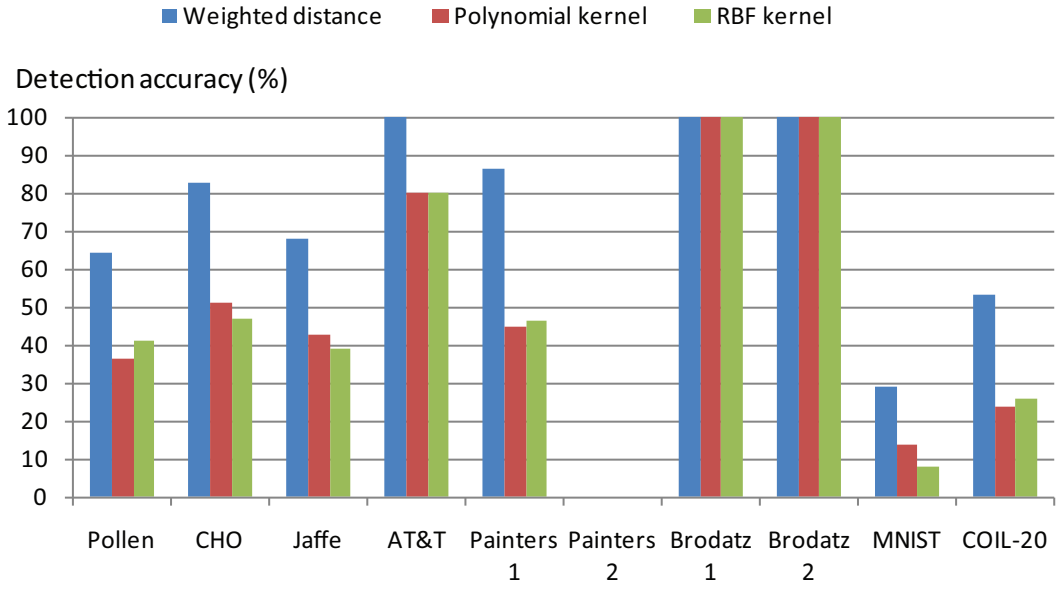


Figure 13. Detection accuracy using the proposed method and one-class SVM using the image feature set.

work efficiently when the variance in the informativeness of the different image features is large. These results are in agreement with previous experiments of automatic image classification using the large image feature set used in this study (Shamir et al., 2010), which also indicated that SVM classifiers have difficulty to effectively handle the strong variance in the informativeness of the image features included in the large feature set. The significant effect of assigning weights to the features compared to using a non-weighted feature space is also discussed in (Orlov et al., 2008;

Shamir et al., 2010).

6. Conclusions

This paper describes a method that applies multi-order image transforms to unsupervised detection of peculiar images in image datasets. The detection of the peculiar images is done in an unsupervised fashion, without prior knowledge that can be used to define the peculiarity of an image in the context of the given image analysis problem at hand. This approach can be useful in cases where it is required to detect unusual images, in the absence of a clear definition of what an unusual image is or how a "peculiar" image is different from a "typical" image in the dataset. For instance, screens in Cell Biology might result in microscopy images of very many cells, and the researcher might be interested in detecting the irregular and uncommon phenotypes (d'Onofrio & Mango, 1984). In many cases the phenotypes of the highest scientific interest can be "new" types of cells, which the researcher has never seen before, and therefore cannot characterize or use previous samples to train a machine vision system to detect. Other examples can include automatic search for peculiar astronomical objects in image datasets acquired by autonomous sky surveys driven by robotic telescopes, or uncommon ground features in datasets of satellites images of the Earth or other planets. Future work will include the application of the proposed system to practical tasks in biology, astronomy, and remote sensing.

The experiments described in this paper show that the detection accuracy of the peculiar image can in some cases be improved if the parameters are adjusted for a specific dataset. However, the pre-defined parameter settings used in this study demonstrated detection accuracy significantly better than random, and showed that in some cases the rank-10 detection accuracy can be as high as 100%. This shows that image features extracted from multi-order image transforms can be used to automatically detect peculiar images in image datasets without using any prior knowledge about the regular images, but more importantly, without any prior knowledge about the target peculiar images.

One limitation of this method is that since the detection of the peculiar image is done in an unsupervised fashion, the feature representation of the regular images should be similar to each other so that the algorithm can differentiate between them and the peculiar image. That is, the variation among the regular images should be smaller than the difference between the peculiar images and the regular images.

Another downside of the method described in this paper is its relatively high computational complexity. Since no prior knowledge about the images can be used, a large and comprehensive set of image features is computed for each image in order to cover many different aspects of the visual content, and then select the most informative features that can differentiate between a regular and a peculiar image. Computing the full set of image content descriptors and transforms can be a computationally expensive task. For instance, computing the feature set for a single 256×256 image takes ~100 seconds using a 2.6GHZ AMD Opteron with 2 GB of RAM. A more comprehensive analysis of the response-time as a function of the image size is available in (Shamir et al., 2008a).

7. Acknowledgments

This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging, and NSF grant number 1157162.

References

Aggarwal, C.C. and P. Yu (2000). Finding generalized projected clusters in high dimensional spaces. *Proc. ACM Intl. Conf. on Management of Data*, pp. 70–81.

Aggarwal, C.C. and P.S. Yu (2001). Outlier detection for high dimensional data. *ACM SIGMOD* **30** pp. 37–46.

Bilenko, M., S. Basu and R. Mooney (2004). Integrating constraints and metric learning in semi-supervised clustering. *Proceedings of the 21st International Conference on Machine Learning*, pp. 81–88.

Boland, M.V. and R. F. Murphy (2001). A Neural Network Classifier Capable of Recognizing the Patterns of all Major Subcellular Structures in Fluorescence Microscope Images of HeLa Cells. *Bioinformatics* **17**, 1213-1223.

Breunig, M.M., H. P. Kriegel, R. T. Ng and J. Sander (2000). LOF: Identifying density-based local outliers. *ACM SIGMOD* **29**, 93–104.

Brodatz, P. Textures, Dover Pub., New York, NY, 1966.

Carpenter, A.E. (2007) Image-based chemical screening. *Nature Chemical Biology* **3**, 461–465.

d'Onofrio, G. and G. Mango (1984). Automated cytochemistry in acute leukemias. *Acta Haematologica* **72**, 221–230.

Duller., A.W.G., G.A.T Duller, I. France and H.F. Lamb (1999). A pollen image database for evaluation of automated identification systems. *Quaternary Newsletter* **89**, 4–9.

Fan, H., O.R. Zaiane, A. Foss and J. Wu (2006). A nonparametric outlier detection for effectively discovering top-N outliers from engineering data. *Proc. Advances in Knowledge Discovery and Data Mining*, pp. 557–566.

Fan, R.E., P.H. Chen and C.J. Lin (2005). Working set selection using the second order information for training SVM. *Journal of Machine Learning Research* **6**, 1889–1918.

Guha, S. (2001). Cure: an efficient clustering algorithm for large databases. *Information Systems* **26**, 35–58.

Gurevich, I.B. and I.V. Koryabkina (2006). Comparative analysis and classification of features for image models. *Pattern Recognition and Image Analysis* **16**, 265–297.

Hodge, V. and J. Austin (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review* **22**, 85–126.

Jones, T.R., A.E. Carpenter, M.R. Lamprecht, J. Moffat, S. J. Silver, J. K. Grenier, A.B. Castoreno, U.S. Eggert, D.E. Root, P. Golland, and D.M. Sabatini (2009). Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Publications of the National Academy of Science* **106**. 1826–1831.

Kameyama, K., S.N. Kim, M. Suzuki, K. Toraichi, and T. Yamamoto (2006). Content-based image retrieval of Kaou images by relaxation matching of region features. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **14**, 509–523.

Knorr E. and R. Ng (1999). Finding intensional knowledge of distance-based outliers. *Proc. VLDB*, pp. 211–222.

LeCun, Y., L. Bottou, Y. Bengio and P. Haffner (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**, 2278–2324.

Liu, C.L., K. Nakashima, H. Sako and H. Fujisawa (2003). Handwritten digit recognition: benchmarking of state-of-the-art techniques. *Pattern Recognition* 36, 2271–2285.

Lukashevich, H., S. Nowak and P. Dunker (2009). Using one-class SVM outliers detection for verification of collaboratively tagged image training sets. *Proc. IEEE ICME*, pp. 682–685.

Lynos, M., S. Akamatsu, M. Kamachi and J. Gyboa (1998). Coding facial expressions with Gabor wavelets. *Proceedings of the Third IEEE International Conference on Automatic Face and Facial Recognition*, 200–205.

Nene, S.A., S.K. Nayar and H. Murase, Columbia Object Image Library (COIL-20), Technical Report No. CUCS-006-96. Columbia University, 1996.

Orlov, N., L. Shamir, T. Macura, J. Johnston, D. M. Eckley and I. G. Goldberg (2008). WND-CHARM: Multi-purpose Image Classification Using Compound Image Transforms. *Pattern Recognition Letters* 29, 1684–1693.

Rodenacker, K. and E. Bengtsson (2003). A feature set for cytometry on digitized microscopic images. *Anal. Cell. Pathol.* **25**, 1–36.

Roth, V. (2005). Outlier detection with one-class kernel Fisher discriminants. *Proc. Advances in Neural Information Processing Systems*, pp. 1169–1176.

Samaria, F. and A. Harter, A. (1994). Parameterisation of a stochastic model for human face identification. *Proc. of the 2nd IEEE Workshop on Applications of Computer Vision*, pp. 138–142.

Scholkopf, B., J. Platt, J. Shawe-Taylor, A.J. Smola and R.C. Williamson. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation* **13**, 1443–1471.

Shamir, L. (2006). Human perception-based color segmentation using fuzzy logic. *International Conference on Image Processing Computer Vision and Pattern Recognition* 2, 496–505.

Shamir, L., N. Orlov, D. M. Eckley, T. Macura, J. Johnston and I. G. Goldberg (2008a). Wndchrm - an open source utility for biological image analysis. *Source Code for Biology and Medicine* 3, 13.

Shamir, L., N. Orlov, D.M. Eckley, T. Macura and I. G. Goldberg (2008b). IICBU 2008 - A proposed benchmark suite for biological image analysis. *Medical & Biological Engineering & Computing* **46**, 943–947.

Shamir, L. (2008) Evaluation of Face Datasets as Tools for Assessing the Performance of Face Recognition Methods. *International Journal of Computer Vision* 79, 225–230.

Shamir, L., S. M. Ling, W. Scott, A. Boss, T., N. Orlov, Macura, D. M. Eckley, L. Ferrucci and I. Goldberg (2009) Knee X-ray image analysis method for automated detection of Osteoarthritis. *IEEE Transactions on Biomedical Engineering* 56, 407–415.

Shamir, L., N. Orlov and I. Goldberg, "Evaluation of the informativeness of multi-order image transforms," *International Conference on Image Processing Computer Vision and Pattern Recognition*, 2009, 37–42.

Shamir, L., T. Macura, N. Orlov, D. M. Eckley and I. G. Goldberg (2010). Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. *ACM Transactions on Applied Perception* **7**, 2:8.

Shamir, L. (2011) Assessing the efficacy of low-level image content descriptors for computer-based fluorescence microscopy image analysis. *J*ournal of Microscopy **243(3)**, 284–292.

Zhang, K., S. Shi, H. Gao and J. Li (2007). Unsupervised outlier detection in sensor networks using aggregation tree. *Lecture Notes in Artificial Intelligence* **4632**, 158–169.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 32–37

On an Unified Class of Functions of Complex Order

T. M. Seoudy[a,*], M. K. Aouf[b]

[a]*Department of Mathematics, Faculty of Science, Fayoum University, Fayoum 63514, Egypt.*
[b]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.*

Abstract

In this paper, we obtain a necessary and sufficient condition for functions in an unified class of functions of complex order. Some of our results generalize previously known results.

Keywords: Analytic function, univalent, starlike, convex, subordination.
2012 MSC: 30C45.

1. Introduction

Let $\mathcal{A}$ denote the class of functions of the form

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n, \tag{1.1}$$

which are analytic in the open unit disk $\mathbb{U} = \{z \in \mathbb{C} : |z| < 1\}$. Suppose that f and g are analytic in $\mathbb{U}$. We say that the function f is subordinate to g in $\mathbb{U}$, or g superordinate to f in $\mathbb{U}$, and we write $f \prec g$ or $f(z) \prec g(z)$ $(z \in \mathbb{U})$, if there exists an analytic function ω in $\mathbb{U}$ with $\omega(0) = 0$ and $|\omega(z)| < 1$, such that $f(z) = g(\omega(z))$ $(z \in \mathbb{U})$. If g is univalent in $\mathbb{U}$, then the following equivalence relationship holds true, see (Miller & Mocanu, 1981) and (Miller & Mocanu, 2000):

$$f(z) \prec g(z) \Longleftrightarrow f(0) = g(0) \quad \text{and} \quad f(\mathbb{U}) \subset g(\mathbb{U}).$$

Let $\mathcal{S}$ be the subclass of $\mathcal{A}$ consisting of univalent functions. Let $\phi(z)$ be an analytic function with positive real part on ϕ with $\phi(0) = 1$, $\phi'(0) > 0$ which maps the unit disk $\mathbb{U}$ onto a region starlike with respect to 1 which is symmetric with respect to the real axis. Let $\mathcal{S}^*(\phi)$ be the class of functions in $f \in \mathcal{S}$ for which

$$\frac{zf'(z)}{f(z)} \prec \phi(z), \tag{1.2}$$

*Corresponding author
Email addresses: tms00@fayoum.edu.eg (T. M. Seoudy), mkaouf127@yahoo.com (M. K. Aouf)

and $C(\phi)$ class of functions in $f \in \mathcal{S}$ for which

$$1 + \frac{zf''(z)}{f'(z)} \prec \phi(z). \tag{1.3}$$

These classes were introduced and studied by (Ma & Minda, 1992). (Ravichandran *et al.*, 2005) defined classes $\mathcal{S}_b^*(\phi)$ and $C_b(\phi)$ of complex order defined as follows :

$$\mathcal{S}^*(\phi; b) = \left\{ f \in \mathcal{A} : 1 + \frac{1}{b}\left(\frac{zf'(z)}{f(z)} - 1\right) \prec \phi(z) \quad (b \in \mathbb{C}^* = \mathbb{C}\setminus\{0\}) \right\} \tag{1.4}$$

and

$$C(\phi; b) = \left\{ f \in \mathcal{A} : 1 + \frac{1}{b}\frac{zf''(z)}{f'(z)} \prec \phi(z) \quad (b \in \mathbb{C}^*) \right\}. \tag{1.5}$$

From (1.4) and (1.5), we have

$$f \in C(\phi; b) \Longleftrightarrow zf' \in \mathcal{S}^*(\phi; b).$$

Now, we introduce a more general class of complex order $\mathcal{T}(\phi; \lambda, b)$ as follows:

Definition 1.1. Let $\phi(z)$ be an analytic function with positive real part on ϕ with $\phi(0) = 1$, $\phi'(0) > 0$ which maps the unit disk $\mathbb{U}$ onto a region starlike with respect to 1 which is symmetric with respect to the real axis. Then the class $\mathcal{T}(\phi; \lambda, b)$ consists of all analytic functions $f \in \mathcal{A}$ satisfying:

$$1 + \frac{1}{b}\left[(1-\lambda)\frac{zf'(z)}{f(z)} + \lambda\left(1 + \frac{zf''(z)}{f'(z)}\right) - 1\right] \prec \phi(z) \ (b \in \mathbb{C}^*; \lambda \geq 0). \tag{1.6}$$

We note that

(*i*) $\mathcal{T}(\phi; 0, b) = \mathcal{S}^*(\phi; b)$ and $\mathcal{T}(\phi; 1, b) = C(\phi; b)$ (Ravichandran *et al.*, 2005);

(*ii*) $\mathcal{T}(\phi; 0, 1) = \mathcal{S}^*(\phi)$ and $\mathcal{T}(\phi; 1, 1) = C(\phi)$ (Ma & Minda, 1992);

(*iii*) $\mathcal{T}\left(\frac{1+(1-2\alpha)z}{1-z}; 0, b\right) = \mathcal{S}_\alpha^*(b)$ and $\mathcal{T}\left(\frac{1+(1-2\alpha)z}{1-z}; 1, b\right) = C_\alpha(b)$ $(0 \leq \alpha < 1; b \in \mathbb{C}^*)$ (Frasin, 2006);

(*iv*) $\mathcal{T}\left(\frac{1+z}{1-z}; 0, b\right) = \mathcal{T}\left(\frac{1+(2b-1)z}{1-z}; 0, 1\right) = \mathcal{S}^*(b)$ $(b \in \mathbb{C}^*)$ (Nasr & Aouf, 1985) and (Wiatrowski, 1970);

(*v*) $\mathcal{T}\left(\frac{1+z}{1-z}; 1, b\right) = \mathcal{T}\left(\frac{1+(2b-1)z}{1-z}; 1, 1\right) = C(b)$ $(b \in \mathbb{C}^*)$ (Nasr & Aouf, 1982) and (Wiatrowski, 1970);

(*vi*) $\mathcal{T}\left(\frac{1+z}{1-z};0,1-\alpha\right)=\mathcal{T}\left(\frac{1+(1-2\alpha)z}{1-z};0,1\right)=\mathcal{S}^{*}(\alpha)$

and $\mathcal{T}\left(\frac{1+z}{1-z};1,1-\alpha\right)=\mathcal{T}\left(\frac{1+(1-2\alpha)z}{1-z};1,1\right)=C(\alpha)\,(0\leq\alpha<1)$ (Robertson, 1936);

(*vii*) $\mathcal{T}\left(\frac{1+z}{1-z};0,be^{-i\gamma}\cos\gamma\right)=\mathcal{S}^{\gamma}(b)$ and $\mathcal{T}\left(\frac{1+z}{1-z};1,be^{-i\gamma}\cos\gamma\right)=C^{\gamma}(b)\left(|\gamma|<\frac{\pi}{2},b\in\mathbb{C}^{*}\right)$ (Al-Oboudi & Haidan, 2000) and (Aouf *et al.*, 2005).

Motivated essentially by the aforementioned works, we obtain certain necessary and sufficient conditions for the unified class of functions $\mathcal{T}(\phi;\lambda,b)$ which we have defined. The motivation of this paper is to generalize the results obtained by (Ravichandran *et al.*, 2005), (Aouf *et al.*, 2005), (Srivastava & Lashin, 2005) and also (Obradovic *et al.*, 1989).

2. Main Results

Unless otherwise mentioned, we assume throughout the sequel that $b\in\mathbb{C}^{*}$, $\lambda\geq 0$ and all powers are understood as principle values. To prove our main result, we need the following lemmas.

Lemma 2.1. *(Ruscheweyh, 1982) Let ϕ be a convex function defined on $\mathbb{U}$, $\phi(0)=1$. Define $F(z)$ by*

$$F(z)=z\,\exp\left(\int_{0}^{z}\frac{\phi(t)-1}{t}dt\right). \tag{2.1}$$

Let $p(z)=1+p_1z+p_2z^2+\ldots$ be analytic in $\mathbb{U}$. Then

$$1+\frac{zq'(z)}{q(z)}\prec\phi(z) \tag{2.2}$$

if and only if for all $|s|\leq 1$ and $|t|\leq 1$, we have

$$\frac{p(tz)}{p(sz)}\prec\frac{sF(tz)}{tF(sz)}. \tag{2.3}$$

Lemma 2.2. *(Miller & Mocanu, 2000) Let $q(z)$ be univalent in $\mathbb{U}$ and let $\varphi(z)$ be analytic in a domain containing $q(\mathbb{U})$. If $\frac{zq'(z)}{q(z)}$ is starlike, then*

$$zp'(z)\varphi(p(z))\prec zq'(z)\varphi(q(z)),$$

then $p(z)\prec q(z)$ and $q(z)$ is the best dominant.

Theorem 2.1. *Let $\phi(z)$ and $F(z)$ be as in Lemma 2.1. The function $f\in\mathcal{T}(\phi;\lambda,b)$ if and only if for all $|s|\leq 1$ and $|t|\leq 1$, we have*

$$\left[\left(\frac{sf(tz)}{tf(sz)}\right)^{1-\lambda}\left(\frac{f'(tz)}{f'(sz)}\right)^{\lambda}\right]^{\frac{1}{b}}\prec\frac{sF(tz)}{tF(sz)}. \tag{2.4}$$

Proof. Define the function $p(z)$ by

$$p(z) = \left[\frac{f(z)}{z}\left(\frac{zf'(z)}{f(z)}\right)^{\lambda}\right]^{\frac{1}{b}} \quad (z \in \mathbb{U}). \tag{2.5}$$

Taking logarithmic derivative of (2.5), we get

$$1 + \frac{zp'(z)}{p(z)} = 1 + \frac{1}{b}\left[(1-\lambda)\frac{zf'(z)}{f(z)} + \lambda\left(1 + \frac{zf''(z)}{f'(z)}\right) - 1\right].$$

Since $f \in \mathcal{T}(\phi; \lambda, b)$, then we have

$$1 + \frac{zp'(z)}{p(z)} \prec \phi(z)$$

and the result now follows from Lemma 2.1. □

Putting $\lambda = 0$ in Theorem 2.1, we obtain the following result of (Shanmugam *et al.*, 2009).

Corollary 2.1. *Let $\phi(z)$ and $F(z)$ be as in Lemma 2.1. The function $f \in \mathcal{S}^*(\phi; b)$ if and only if for all $|s| \leq 1$ and $|t| \leq 1$, we have*

$$\left(\frac{sf(tz)}{tf(sz)}\right)^{\frac{1}{b}} \prec \frac{sF(tz)}{tF(sz)}. \tag{2.6}$$

For $\lambda = 1$ in Theorem 2.1, we obtain the following result of (Shanmugam *et al.*, 2009).

Corollary 2.2. *Let $\phi(z)$ and $F(z)$ be as in Lemma 2.1. The function $f \in C(\phi; b)$ if and only if for all $|s| \leq 1$ and $|t| \leq 1$, we have*

$$\left(\frac{f'(tz)}{f'(sz)}\right)^{\frac{1}{b}} \prec \frac{sF(tz)}{tF(sz)}. \tag{2.7}$$

Theorem 2.2. *Let $\phi(z)$ be starlike with respect to 1 and $F(z)$ given by (2.1) be starlike. If $f \in \mathcal{T}(\phi; \lambda, b)$, then we have*

$$\frac{f(z)}{z}\left(\frac{zf'(z)}{f(z)}\right)^{\lambda} \prec \left(\frac{F(z)}{z}\right)^{b}. \tag{2.8}$$

Proof. Let $p(z)$ be given by (2.5) and $q(z)$ be given by

$$q(z) = \frac{F(z)}{z} \quad (z \in \mathbb{U}). \tag{2.9}$$

After a simple computation we obtain

$$1 + \frac{zp'(z)}{p(z)} = 1 + \frac{1}{b}\left[(1-\lambda)\frac{zf'(z)}{f(z)} + \lambda\left(1 + \frac{zf''(z)}{zf'(z)}\right) - 1\right].$$

and

$$\frac{zq'(z)}{q(z)} = \frac{zF'(z)}{F(z)} - 1 = \phi(z) - 1.$$

Since $f \in \mathcal{T}(\phi; \lambda, b)$, we have

$$\frac{zp'(z)}{p(z)} \prec \frac{zq'(z)}{q(z)}.$$

The result now follows by an application of Lemma 2.2. □

Putting $\lambda = 0$ in Theorem 2.2, we obtain the following results of (Shanmugam *et al.*, 2009).

Corollary 2.3. *Let $\phi(z)$ be starlike with respect to 1 and $F(z)$ given by* (2.1) *be starlike. If $f \in \mathcal{S}^*(\phi; b)$, then we have*

$$\frac{f(z)}{z} \prec \left(\frac{F(z)}{z}\right)^b.$$

Taking $\phi(z) = \dfrac{1+Az}{1+Bz}$ $(-1 \le B < A \le 1)$ in Theorem 2.2, we get the following corollary.

Corollary 2.4. *If $f \in \mathcal{T}\left(\dfrac{1+Az}{1+Bz}; \lambda, b\right)(-1 \le B < A \le 1)$, then we have*

$$\frac{f(z)}{z}\left(\frac{zf'(z)}{f(z)}\right)^{\lambda} \prec (1+Bz)^{\frac{(A-B)b}{B}} \quad (B \neq 0).$$

For $\phi(z) = \dfrac{1+z}{1-z}$ and $\lambda = 0$ in Theorem 2.2, we get the following result of (Obradovic *et al.*, 1989), and (Srivastava & Lashin, 2005).

Corollary 2.5. *If $f \in \mathcal{S}^*(b)$, then we have*

$$\frac{f(z)}{z} \prec (1-z)^{-2b}.$$

Putting $\phi(z) = \dfrac{1+z}{1-z}$ and $\lambda = 1$ in Theorem 2.2, we get the following result of (Obradovic *et al.*, 1989), and (Srivastava & Lashin, 2005).

Corollary 2.6. *If $f \in C(b)$, then we have*

$$f'(z) \prec (1-z)^{-2b}.$$

For $\phi(z) = \dfrac{1+z}{1-z}, \lambda = 0$ and replacing b by $be^{-i\gamma}\cos\gamma$ $\left(|\gamma| < \dfrac{\pi}{2}, b \in \mathbb{C}^*\right)$ in Theorem 2.2, we get the following result of (Aouf *et al.*, 2005).

Corollary 2.7. *If $f \in \mathcal{S}^{\gamma}(b)$, then we have*

$$\frac{f(z)}{z} \prec (1-z)^{-2be^{-i\gamma}\cos\gamma} .$$

Taking $\phi(z) = \dfrac{1+z}{1-z}$, $\lambda = 1$ and replacing b by $be^{-i\gamma}\cos\gamma \left(|\gamma| < \dfrac{\pi}{2}, b \in \mathbb{C}^*\right)$ in Theorem 2.2, we get the following result of (Aouf *et al.*, 2005).

Corollary 2.8. *If $f \in C^{\gamma}(b)$, then we have*

$$f'(z) \prec (1-z)^{-2be^{-i\gamma}\cos\gamma} .$$

References

Al-Oboudi, F. M. and M. M. Haidan (2000). Spirallike functions of complex order. *J. Natural Geom.* **19**, 53–72.

Aouf, M. K., F. M. Al-Oboudi and M. M. Haidan (2005). On some results for λ−spirallike and λ−robertson functions of complex. *Publ. Instit. Math. Belgrade* **77**(91), 93–98.

Frasin, B. A. (2006). Family of analytic functions of complex order. *Acta Math. Acad. Paedagog. Nyházi. (N. S.)* **22**(2), 179–191.

Ma, W. C. and D. Minda (1992). A unified treatment of some special classes of univalent functions. In: *Proceedings of the Conference on Complex Analysis (Tianjin, 1992), Internat. Press, Cambridge, MA*. pp. 157–169.

Miller, S. S. and P. T. Mocanu (1981). Differential subordinations and univalent functions. *Michigan Math. J.* **28**, 157–171.

Miller, S. S. and P. T. Mocanu (2000). *Differential Subordinations: Theory and Applications, Serieson Monographs and Textbooks in Pure and Applied Mathematics*. Vol. 225, Marcel Dekker, New York and Basel.

Nasr, M.A. and M. K. Aouf (1982). On convex functions of complex order. *Mansoura Bull. Sci.* **8**, 565–582.

Nasr, M.A. and M. K. Aouf (1985). Starlike function of complex order. *J. Natur. Sci. Math.* **25**, 1–12.

Obradovic, M., M. K. Aouf and S. Owa (1989). On some results for starlike functions of complex order. *Publ. Inst. Math., Nouv. Ser.* **46**(60), 79–85.

Ravichandran, V., Y. Polatoglu, M. Bolcal and A. Sen (2005). certain subclasses of starlike and convex functions of complex order. *Hacettepe J. Math. Stat.* **34**, 9–15.

Robertson, M. S. (1936). On the theory of univalent functions. *Ann. Math.* **37**, 374–408.

Ruscheweyh, St. (1982). *Convolutions in geometric function theory*. Séminaire de mathématiques supérieures. Montréal, Québec, Canada : Presses de l'Université de Montréal.

Shanmugam, T. N., S. Sivasubramanian and S. Kavitha (2009). On certain subclasses of functions of complex order. *South. Asian Bull. Math.* **33**, 535–541.

Srivastava, H. M. and A.Y. Lashin (2005). Some applications of the briot-bouquet differential subordination. *J. Inequal. Pure Appl. Math.* **6**(2, Art. 41), 1–7.

Wiatrowski, P. (1970). On the coefficients of some family of holomorphic functions. *Zeszyty Nauk. Uniw. Łódz Nauk. Mat.-Przyrod.* **39**, 75–85.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 38–55

The Scale-Curvature Connection and its Application to Texture Segmentation

Eli Appleboim[a], Yedidya Hyams[a], Shai Krakovski[a], Chen Sagiv[b], Emil Saucan[c,*]

[a]*Department of Electrical Engineering, Technion, Haifa 32000, Israel*
[b]*SagivTech Ltd., Eliezer Yaffe 37/18, Ra'anana 43451, Israel*
[c]*Department of Mathematics, Technion, Haifa 32000, and Department of Mathematics and Computer Science, The Open University of Israel, Ra'anana 43537, Israel*

Abstract

In this work we establish a theoretical relation between the notions of scale and a discrete Finsler-Haantjes curvature. Based on this connection we demonstrate the applicability of the interpretation of scale in terms of curvature, to signal processing in the context of analysis and segmentation of textures in images. The outcome of this procedure is a novel scheme for texture segmentation that is based on scaled metric curvature. The presented method proves itself to be efficient even when the multiscale analysis is done up to scales of 19 and more. Our main conclusions are that the discrete curvature calculated on sampled images can give us an indication on the local scale within the image, and therefore can be used for many additional tasks in image analysis.

Keywords: Wavelets, scale-space, Finsler-Haantjes curvature, texture segmentation.
2010 MSC: 42C40, 68U10, 51K10.

1. Introduction

Several tasks in image and signal processing require the calculation and usage of scale. Determining the typical scale at some image location can be useful for de-convolution, detection and recognition tasks. The popular image registration algorithms, SIFT (Lowe, 1999) and SURF (Bay *et al.*, 2006) account for the scale at image locations as a pre-processing step for calculating scale invariant key points, that are used in turn for matching.

Typical approaches to calculate scale in the signal processing community rely on analyzing a multi scale representation of the images, via the scale space approach or the wavelet transform.

*Corresponding author
Email addresses: eliap@ee.technion.ac.il (Eli Appleboim), gwavelet gwavelet@gmail.com (Yedidya Hyams), gwavelet gwavelet@gmail.com (Shai Krakovski), chen@sagivtech.com (Chen Sagiv), semil@tx.technion.ac.il (Emil Saucan)

The versatility and adaptability of scale space theory and wavelets for a variety of tasks in Image Processing and related fields is too well established in the scientific community, and the bibliography pertaining to it is far too extensive, to even begin to review it here.

On the other hand, the concept of curvature is well established in the field of computational geometry. Intuitively, scale and curvature are related. High curvature account for phenomenon that happen at smaller scales than those that are related to low curvature. This relation is further stressed analytically and formally in the smooth category, as the curvature of a smooth curve at some point is defined to be the inverse of the squared radius of the osculatory circle at that point so specifically making curvature a function of scale. Curvature decreases as the inverse of the square root of scale (Petersen, 1998).

The multiresolution property of wavelets has been already applied in determining the curvature of planar curves (Antoine & Jaques, 2003) and to the intelligence and reconstruction of meshed surfaces (see, e.g. (Lounsbery *et al.*, 1997), (Valette & Prost, 2004), amongst many others). Moreover, the intimate relation between scale and differentiability in natural images has also been stressed (Florack *et al.*, 1992).

An intriguing issue is whether one can replace the intuitive trade-off between scale and curvature, by a formal concept of *wavelet curvature*, in particular in cases such as the Strömberg wavelets (Strömberg, 1983) that are based on piecewise-linear functions, and if so then,to what extent this can be further extended to the more difficult case of Haar wavelets that are not even piecewise linear and to what extent this can be made general.

Apparently, this can be done by using *metric curvatures* (Blumenthal & Menger, 1970) (and (Saucan, 2006) for a short presentation). It turns out that the best candidate, for the desired metric curvature is the *Finsler-Haantjes curvature*, due to its adaptability to both continuous and discrete settings (see, e.g. (Saucan & Appleboim, 2005), (Saucan & Appleboim, 2009)).

We have first introduced a formal relation between discrete curvature and scale in (Saucan *et al.*, 2010). In the present paper, that represents a continuation of our previous, above mentioned article, we suggest that a simple curvature calculation can replace the tedious work of convolving images with a large number of multi scaled filters. We show how scale and curvature are related to each other, for a variety of families of wavelets. Afterwards we present the Finsler-Haantjes curvature measure for images and develop a novel scheme for texture separation. Our main goal is, however, more far-reaching, namely to try and bridge, at least partly, the the gap between the two basic, largely non-intersecting, approaches prevalent in Image Processing and related fields: The geometric one, that is closely related to the Graphics community philosophy; and the more classical, Fourier Analysis/Wavelets driven one.

The paper is organized as follows: First, we introduce the mathematical background needed and discuss the notion of scale in Section 2.1. We then elaborate on the Finsler-Haantjes curvature in Section 2.2 and introduce the Fisler-Haantjes curvature of wavelets and for images in Section 3. In section 4 we suggest a scheme for texture analysis in images that is based on our discrete scaled curvature measure. In addition to texture segmentation there is a huge variety of further possible applications to the ideas and methods presented herein, as well as open issues for further research. Some of those are mentioned in Section 5 in which we summarize the paper.

2. Mathematical Background

In this section we present both the notion of scale and that of the Finsler-Haantjes Curvature. While these two components are derived from completely different worlds, we show that they are strongly related to each other.

2.1. The Notion of Scale

The notion of scale is fundamental in many mathematical and applicative discussions. Scale is one of these terms that has a clear intuitive meaning, but is hard to be defined mathematically. The question of finding a measure for calculating the local scale in signals and images has been addressed in the past in the context of scale space analysis and wavelets transform. It plays a significant role in the framework of image matching and registration, where scale invariant descriptors are desired. Evaluating the dominant scale within image data is highly important for real life applications. For a computer vision system analyzing an unknown scene, there is no way to know a priori what scales are appropriate for describing the interesting structures in the image data. Hence, the only reasonable approach is to consider descriptions at multiple scales in order to be able to capture the unknown scale variations that may occur. Scale-space theory is a formal theory for handling image structures at different scales, by representing an image as a one-parameter family of smoothed images, the scale-space representation. This representation is parameterized by the size of the smoothing kernel used for suppressing fine-scale structures.

In the early eighties Witkin (Witkin, 1983a),(Witkin, 1983b) proposed to consider scale as a continuous parameter and formulated the principal rules of modern scale-space theory relating image structures represented at different scales. Since then, scale-space representation and its properties have been extensively studied and important contributions have been made by Koenderink (Koenderink, 1984), Lindeberg (Lindeberg, 1998) and Florack (Florack *et al.*, 1992). In many cases it is necessary to select locally appropriate scales for further analysis. This need for scale selection originates from the need to process real-world objects that may have different sizes and because the distance between the object and the camera can vary. The seminal work of Lindenberg (Lindeberg, 1998) dealt with the issue of automatic scale selection. The idea was to determine the characteristic scale for which a given function attains an extremum over scales. The name characteristic is somewhat arbitrary as a local structure can exist at a range of scales and within this range there is no preferred scale of perception. However, a scale can be named characteristic, if it conveys more information comparing to other scales. In his work, Lindenberg noted that a highly useful property of scale-space representation is that image representations can be made invariant to scales, by performing automatic local scale selection based on local maxima (or minima) over scales of normalized derivatives.

This work served as the basis for tasks such as blob detection, corner detection, ridge detection and edge detection. Scale space theory is fundamental for detecting invariant features within signals and images that can be used for various tasks such as registration, detection and recognition among others. Multi-scale representation of data is crucial for extracting local features used for determining regions of interest for subsequent detection of scale-invariant interest points for computing image descriptors, most notable are the SIFT (Lowe, 1999) and SURF (Bay *et al.*, 2006) frameworks.

Both the SIFT and the SURF algorithms rely on Scale-space extrema detection, where the first stage of computation searches over all scales and image locations. It is implemented by using a Laplacian-of-Gaussian or a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.

A common practice for scale determination relies on the convolution of the data with a bank of functions that have different scales. The characteristic scale usually corresponds to the local extremum of the convolution results, taken over scales. The characteristic scale is related to structure. The common methodology for finding this exterma values in scale space involves analysis of the behavior of the Laplacian of Gaussian (Bay *et al.*, 2006), Difference of Gaussian (Lowe, 1999) and the Hessian matrix to name a few. There are strong relations between scale-space theory and wavelet theory, although these two notions of multi-scale representation have been developed from somewhat different premises. Wavelets are multi scaled versions of a specific mother function, thus when convolving them with data, one can exploit the scale contents of that data, in a very similar way that the frequency contents of data can be expressed using the Fourier transform. A strong response to a wavelet function with a certain support and scale, suggests that there is significant information at that scale at that image location.

2.2. The Finsler-Haantjes Curvature

The most intuitive definition for curvature is the amount by which a geometric object deviates from being flat, or straight in the case of a line. It is natural to define the curvature of a straight line to be identically zero. The curvature of a circle of radius R should be large if R is small and small if R is large. Thus the curvature of a circle is defined to be the reciprocal of the squared radius (do Carmo, 1976).

The following metric definition for curvature is due to Haantjes, following an idea of Finsler (Blumenthal & Menger, 1970):

Definition 2.1. *Let (M,d) be a metric space, let* $c : I = [0,1] \tilde{\to} M$ *be a homeomorphism, and let* $p, q, r \in c(I),\ q, r \neq p$. *Denote by* $\widehat{qr}$ *the arc of* $c(I)$ *between* q *and* r, *and by* qr *the segment from* q *to* r. *We say that* c *has* Finsler-Haantjes curvature $\kappa_{FH}(p)$ *at the point* p *iff:*

$$\kappa_{FH}^2(p) = 24 \lim_{q,r \to p} \frac{l(\widehat{qr}) - d(q,r)}{(d(q,r))^3}\ ; \tag{2.1}$$

where "$l(\widehat{qr})$*" denotes the length, in intrinsic metric induced by* d, *of* $\widehat{qr}$ *– see Figure 1. (Here we assume the curve* $c(I)$ *is* rectifiable, *hence that, in particular, the arc* $\widehat{qr}$ *has finite length.)*

This definition of curvature represents, indeed, a proper adaptation, for an extensive class of curves in quite general metric spaces, of the classical notion of curvature, as proven by the following

Theorem 2.1. *Let* $c \in C^3(I)$ *be a smooth curve in* $\mathbb{R}^3$, *and let* $p \in c$ *be a regular point. Then* $\kappa_{FH}(p)$ *exists and, moreover,* $\kappa_{FH}(p) = k(p)$ *– the classical (differential) curvature of* c *at* p.

Remark. Originally the Finsler-Haantjes curvature is defined with $l(\widehat{qr})$ appearing in the denominator instead of $d(q,r)$, ((Blumenthal & Menger, 1970)). We have opted for the above definition

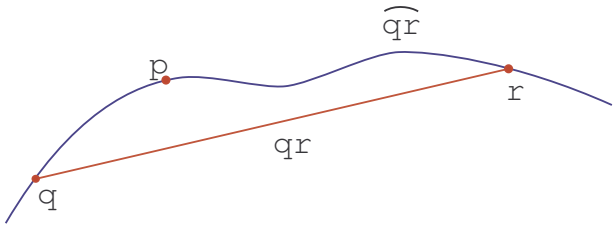


Figure 1. A (metric) arc and its corresponding chord (metric segment).

for practicality reasons. Moreover, in the setting of this work (and, in fact, in a much more general context) the above theorem still holds with our modified definition, therefore the definition used herein is interchangeable with the original one (see, again (Blumenthal & Menger, 1970)).

3. Finsler-Haantjes Curvature for Wavelets and Images

In this section we consider a semi-discrete (or semi-continuous) version of the Finsler-Haantjes curvature, and then introduce this curvature measure in the case of a wavelet function.

3.1. Semi-discrete Finsler-Haantjes Curvature

Consider a typical piecewise-linear wavelet φ, such as the one depicted in Figure 2, let $\widehat{AE}$ be the arc of curve between the points A and E, and let $d(A, E)$ is the length of the line-segment AE.

Then

$$l(\widehat{AE}) = a + b + c + d \ \ ; d(A, E) = e + f. \tag{3.1}$$

The following discretization of formula (2.1) is, therefore, natural:

$$\kappa^2_{FH}(\varphi) = 24[(a + b + c + d) - (e + f)]/(e + f)^3. \tag{3.2}$$

In addition to the total curvature of the wavelet φ, one can also compute the "local" curvatures of the partial wavelets $\varphi_1 = \widehat{ABC}$ and $\varphi_2 = \widehat{CDE}$, that is the curvatures at the "peaks" B and D:

$$\kappa^2_{FH}(B) = 24(a + b - e)/e^3 \ , \tag{3.3}$$

and

$$\kappa^2_{FH}(D) = 24(c + d - f)/f^3 \ , \tag{3.4}$$

as well as the mean curvature of these peaks:

$$H_{FH}(\widehat{AE}) = [\kappa_{FH}(B) + \kappa_{FH}(D)]/2. \tag{3.5}$$

Even though these variations may prove to be useful in certain applications, we believe that the correct approach, in the sense that it best corresponds to the scale of the wavelet, would be to compute the total curvature of φ. However, had the definition of Finsler-Haantjes curvature been limited solely to piecewise-linear wavelets, its applicability would have also been diminished. We show, however, that it is also definable for the "classical" Haar wavelets, in a rather straightforward manner.

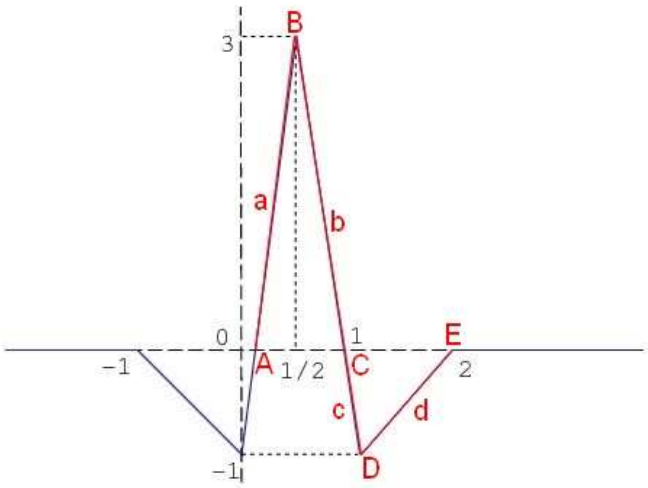


Figure 2. A typical piecewise-linear wavelet (red), part of the Meyer Wavelet (Meyer, 1993) (blue and red).

Remark. In the sequel we will therefore omit the coefficient $\sqrt{24}$ for convenience.

3.2. *Finsler-Haantjes curvature of Haar Wavelets*

For every $s \in \mathbb{Z}$ let $j = 2^s$, and let Ψ_j denote the Haar wavelet at scale j and with zero shift, where $\Psi_1 = \Psi$ is the mother wavelet of Haar basis, considered in the above example. Then Ψ_j can be presented as:

$$\Psi_j = \begin{cases} j^{-1}, & x \in (0, \frac{j}{2}); \\ -j^{-1}, & x \in (\frac{j}{2}, j); \\ 0 & otherwise. \end{cases} \tag{3.6}$$

Then, in the notations of Figure 2 we have that $A = 0, E = j$, so we have that $l(\widehat{AE}) = 4 \cdot j^{-1} + j$, and $d(\widehat{AE}) = j$, therefore for these wavelets the Finsler-Haantjes curvature satisfies:

$$\kappa^2_{FH}(\Psi_j) = \frac{(4 \cdot j^{-1} + j) - j}{j^3} = 4 \cdot j^{-4} \tag{3.7}$$

The Finsler-Haantjes curvature is certainly invariant under shifts therefore the same dependency of K_{FH} in scale is the same for shifted Haar wavelets as well.

3.3. *Finsler-Haantjes curvature of Walsh Basis*

Let R_s be the Rademacher function which takes the value $1, -1$ on the dyadic intervals $[\frac{j}{2^{s+1}}, \frac{j+1}{2^{s+1}})$, $j = 0, 1, ..., 2^{s+1}$ of the unit interval. Figure 3 shows the first four Rademacher functions.

Then for any $k \in \mathbb{N}$, if we take the binary expansion of k as a sum of powers of 2, $k = 2^{p_1} + 2^{p_2} + \cdots + 2^{p_m}$, and define the k_{th} Walsh function as ((Beauchamp, 1975)):

$$W_k = R_{p_1} \cdot R_{p_2} \cdots R_{p_m} \tag{3.8}$$

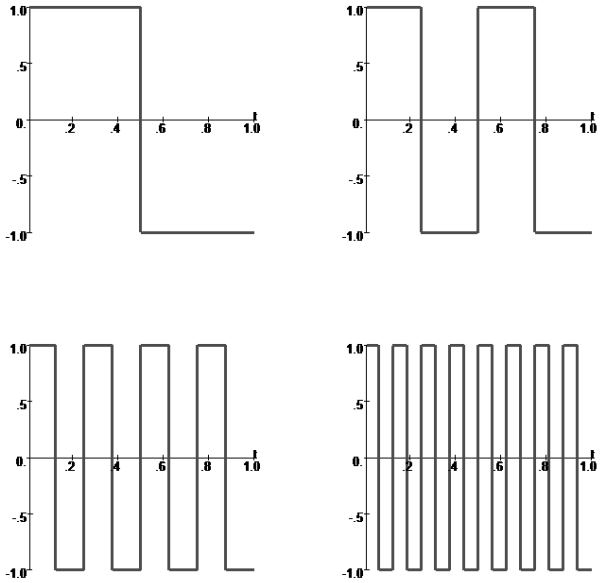


Figure 3. First four Rademacher functions.

Again, in the notations of Figure 2 we have that, $A = 0, E = 1, l(\widehat{AE}) = 2 \cdot 2^s + 1, d(\widehat{AE}) = 1$ which results with

$$\kappa^2_{FH}(W_s) = \frac{(2 \cdot 2^s + 1) - 1}{1^3} = 2^{s+1} \ . \tag{3.9}$$

Although the Walsh basis is not a wavelet basis (see (Beauchamp, 1975)), we can easily regard the function W_s as a function in a specific scale which is $j = \frac{1}{s+1}$ hence the curvature of the Walsh function at scale j is $2^{\frac{1}{j}}$.

For a smooth wavelet Ψ, compactly supported, we can of course define its Finsler-Haantjes, K_{FH} by taking $l(\widehat{AE})$ to be the usual arc length given by $\int_{\text{support}\Psi} \sqrt{1 + \Psi'^2}$.

3.4. Curvature vs. Scale

By an analysis similar to those in Sections 3.2 and 3.3, we will be able to compute a specific dependency of the Finsler-Haantjes curvature as a function of scale, for every family of piecewise constant wavelets. Moreover, with only a limited amount of additional effort, this goal is also achievable for the families of piecewise linear wavelets. This correlation most probably depends on the specific family under consideration. Indeed, as already noted above, Haar wavelets behave differently from the Walsh basis as far as the curvature vs. scale correspondence is concerned. To obtain a similar relation in the case of smooth wavelets, one should recall that if Ψ_j is a smooth wavelet function at some scale j, then it can be approximated, for instance, in the L_2-norm, by a sequence of Haar functions. More precisely, for every $\epsilon > 0$, there exists $k = k(j, \epsilon) \in \mathbb{N}$, such that

$$\int_{\text{supp}\Psi_j} |\Psi_j - \sum_i^k Haar_i|^2 < \epsilon \ . \tag{3.10}$$

Combining the inequality above with Equation (3.7) we obtain that the Finsler-Haantjes curvature of Ψ will display a scale-curvature relation similar with that observed for the Haar wavelet. Here, by "similar" we mean that it will decrease as curvature increases. However, we probably cannot expect in this case a simple analytic expression comparable with the one displayed in the case of Haar wavelets. The precise behavior, if can be derived at all, is left as an open question at this point. It would be reasonable to assume that it depends on the specific wavelet family Ψ, as well as on the proximity factor ϵ. Nevertheless, this behavior is indeed demonstrated in the numerical tests that were applied on a variety of wavelets families. For each type of wavelets, Haantjes curvature was computed for the wavelets families at different scales. The general behavior is similar in all different families and are shown in Figure 4. It is shown that, as expected, curvature decreases as scale increases. As indicated by Equation (3.7), we see that the decrease of curvature as a function of scale for the Haar wavelet is different then the one of classical differential geometry of smooth manifolds, where the decrease has a magnitude of $(scale)^{-2}$, while for the Haar wavelet it was shown to be in a magnitude of $(scale)^{-4}$. The difference evidently follows from the fact that we compute the Finsler-Haantjes curvature in a global way rather than locally, as usually curvature is computed in the classical differential geometric setting.

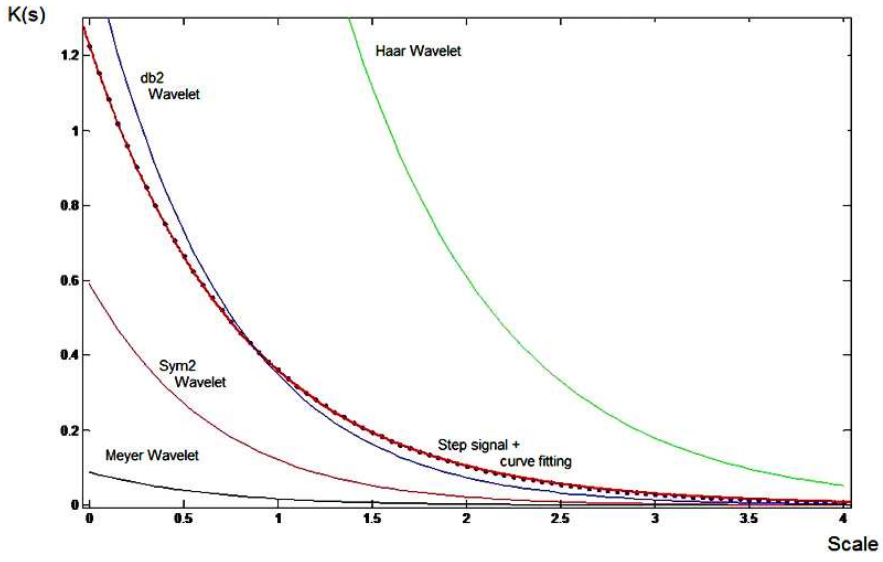


Figure 4. Curvature as a function of scale for a number of standard wavelets: The Haar wavelets, db2 wavelets, sym2 wavelets, the Meyer wavelets; as well as the step signal.

3.5. Curvature for images

From the definition of Finsler-Haatjes for curves we can easily define a discrete version of curvature for surfaces in general, and for images in particular. For say, a point x on a surface Σ, the most natural thing to do is to consider Fisler-Haantjes curvature in any of the directions emanating from x, then find the maximal and minimal curvatures, and then take either the mean of these two

values so to obtain a Finsler-Haantjes *mean curvature* or, alternatively take the multiplication of these two in order to obtain a Finsler-Haantjes version of the *Gaussian curvature* at x. We adopt this concept to images while we consider an image as a surface embedded in some $\mathbb{R}^n$. A gray scale image, for example, can naturally be considered as a surface in $\mathbb{R}^3$. In this case the Finsler-Haantjes curvature is computed at each pixel in four different directions and then the average of these four curvatures is taken. This is done in the framework that was defined in Equation 3.3 of localized curvature, where for images the localization is done by considering a window of some size $n \times n$ centered at the pixel. Each of the images shows the results of computing this version of mean curvature as computed for window sizes of 3×3, 5×5 and 7×7.

Before we proceed further, let us briefly discuss a limit case: If a signal (image) displays a unique scale, e.g. for periodic signals, for which there is a direct correlation between scale and the period T, one expects that to observe that the graph of the curvature function is the smoothest precisely in windows of size T. That this is indeed the case is illustrated in Figure 5.

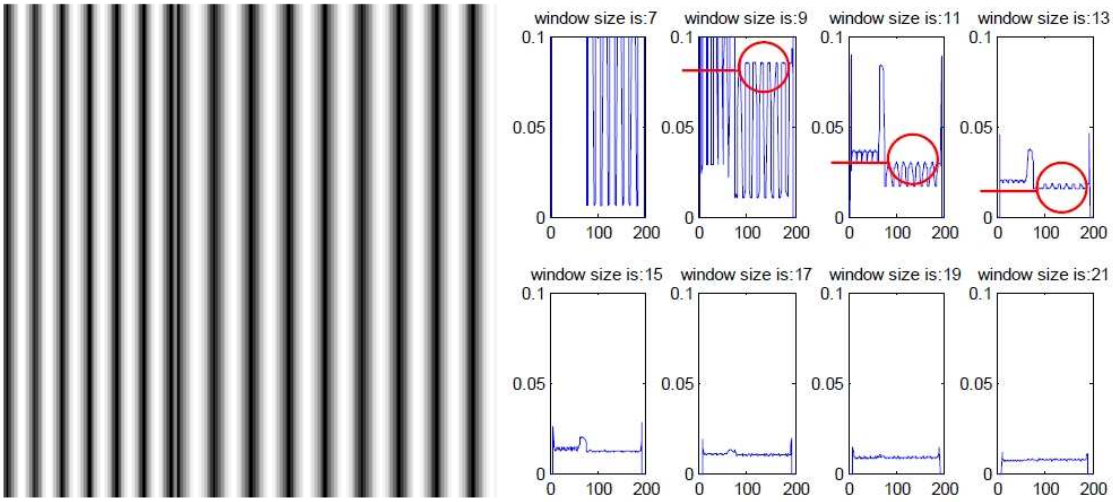


Figure 5. Left: A test image, consisting of two *sin* signals of different periods (11 pixels – left, and 15 pixels – right). Right: As expected, on the windows corresponding to the period of the signal (image) the curvature graph is the most smooth. (Notice the highlighted "windows".)

We clearly can see how Haanjes-Finsler curvature performs as an edge detector. This result is expected, since curvature, even its metric, abstract setting, still has qualities similar to those of a second derivative. This is clearly illustrated in Figure 6. This is further emphasized in the more challenging example in Figure 7, of a satellite image of the Egypt pyramids at Giza, as they lie against the background of sand dunes and opposed to the adjacent neighbourhood of Cairo. It should be noted that, curvature map in itself can serve as a good man made detection tool in arial and satellite images.

Figure 6. Haantjes-Finsler curvature of an image with respect to different scales. From top left in clockwise direction: original image, $3X3$, $7X7$ and $5X5$ window size.

4. Texture segmentation

As a possible application of the proposed method of indicating scale via curvature we look at the problem of image texture segmentation. The novel segmentation scheme yielded from this approach is outlined below.

1. At each pixel Haantjes-Finsler curvatures are computed at different scales and different orientations. For each window, curvatures are computed in 4 directions, horizontal vertical and two diagonal directions $\kappa_h, \kappa_v, \kappa_{d_1}, \kappa_{d_2}$. Finally, the average $\kappa_{Avg} = (\kappa_{Max}(pix) - \kappa_{min}(pix))/\kappa_{Max}(pix)$ of these four obtained curvatures is taken as the curvature at the pixel in the relevant scale. (The specific average considered here was inspired by the standard Image Processing definition of the contrast $C(I)$ of an image I, $C(I) = (I_{Max} - I_{min})/I_{Max}$.) This approximates the average curvature at each scale. The outcome of this step is a vector of length m where m is the number of different scales, $V(pix, scale)$ where pix denotes pixel, and each entry of the vector represents the average curvature at the corresponding scale.

2. Next, at each pixel, the gradient, with respect to scale $\nabla_{scale}V$, of this curvature vector is computed, and we look for the scales at which the gradient exceeds a predefined threshold. (Note that, at this point, curvature and scale are already interchangeable.) Afterwards, all scales which fulfill the threshold criteria are averaged in order to get a scalar value for each pix. The average scale is the outcome of this step. We consider this scale as the scale of important information at the relevant pixel.

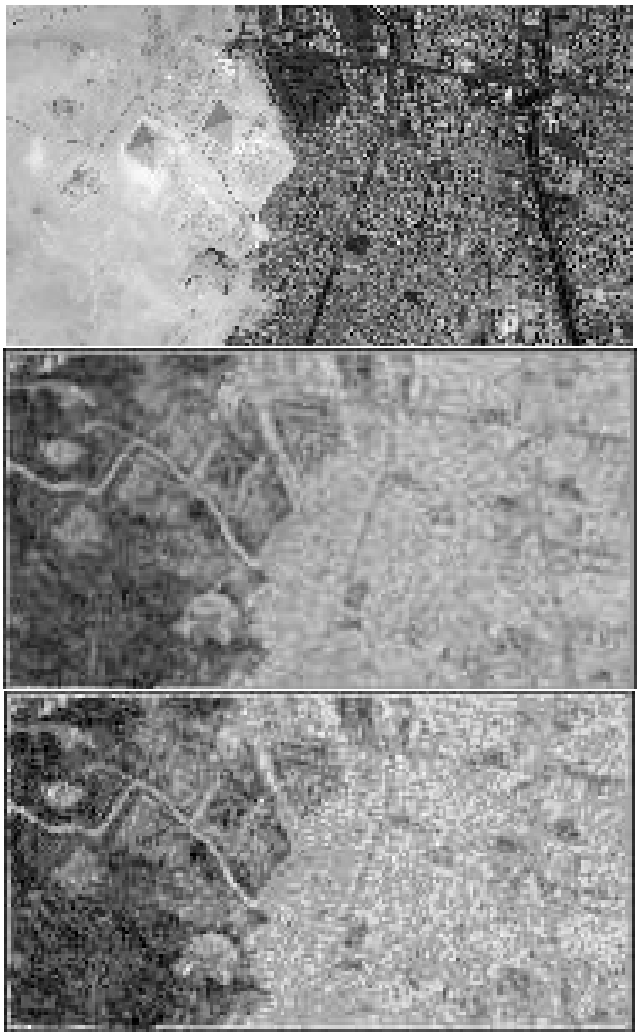


Figure 7. Haantjes-Finsler curvature of a satellite image with respect to different scales. From top to bottom: original image, curvature averaged on 3X3 windows, curvature averaged on 7X7 windows.

3. The output of previous step is a matrix in the same size as the image size, each entry of which is the scale of information at the relevant pixel. A smoothed version of this matrix is obtained by a linear filtering at size which is compatible with the amount of localized information one wishes to obtain. Segmentation to small textures will require small filtering support.

4. We segment the image according to the smoothed information scale computed in the previous step. Pixels with similar scales are grouped together to form different segments. The segmentation is done after curvature values are quantized to several levels. In the experiments shown herein quantization is taken into seven levels.

The procedure detailed above is summarized as Algorithm 1, which is divided into its four main constituent parts:

Input: Grayscale image I
Output: Vector $V(pix, scale)$ of length m = number of scales
foreach *pixel pix in I* **do**
  foreach *window of size* $\leq m$ **do**
    compute $\kappa_h(pix), \kappa_v(pix), \kappa_{d_1}(pix), \kappa_{d_2}(pix)$ and find $\kappa_{Max}(pix), \kappa_{min}(pix)$;
    compute $\kappa(pix) = \kappa_{Avg}(pix) = \frac{\kappa_{Max} - \kappa_{min}}{\kappa_{Max}}$;
  end
end

Input: $V(pix, scale)$
Output: Matrix $M(I)$ – The average scale matrix
foreach *pixel pix in I* **do**
  compute $\nabla_{scale} V$;
  choose scale threshold s_0;
  select scales s_i for which $V(pix, scale) > s_0$;
  compute $s_{Avg} = s_{Avg}(pix) = \frac{\sum s_i}{|\{s_i \,|\, V(pix,scale) > s_0\}|}$;
end

Input: $M(I)$
Output: Matrix $\widetilde{M}(I)$ – Smoothed version of $M(I)$
choose window size $w_0 = w_0(texture)$;
apply linear filter at size w_0;

Input: $\widetilde{M}(I)$
Output: Segmented image $\tilde{I}$
choose maximal number of quantization levels q_0;
foreach *pixel pix in I* **do**
  compute the quantized values $\bar{s}_{Avg}(pix)_j\,,\ j = 1, \ldots, q_0$;
end
group pixels in segments $\mathfrak{s}_l(j) = \{pix \,|\, \bar{s}_{Avg}(pix) = \bar{s}_{Avg}(pix)_j\}$;

Algorithm 1: Segmentation algorithm.

In the following figures, first results of the proposed method are shown for different images. Figure 8 shows the original synthetic image which is comprised of two different textures, the second image in the figure shows the gradient vector $\frac{\partial}{\partial scale} V(scale, pixel)$ of the scaled curvature vector V, while in the third image we see the gradient vector field after smoothing with a filter of size 3×3, and in the fourth image the outcome segmentation is shown after quantization of the smoothed gradient into 4 levels. In the segmented image we see intermediate texture around the internal circle. When one looks at the original image we can clearly see that this is caused by those pixels that are in the intersection of the two different ares of the images and indeed we cannot associate specific texture to these pixels.

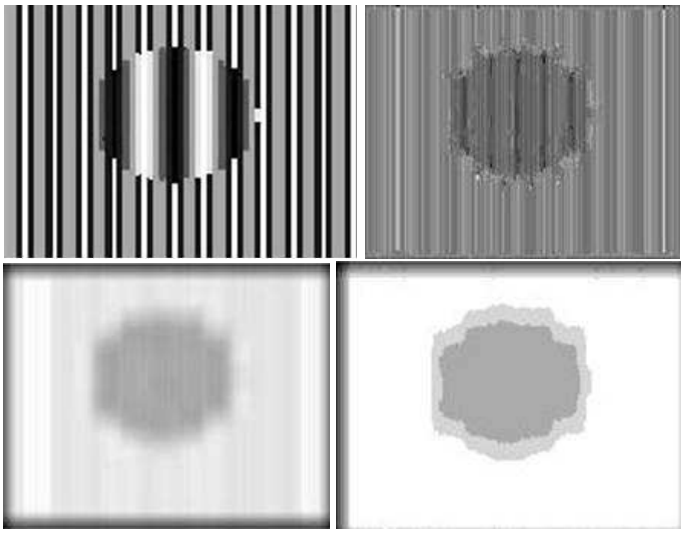


Figure 8. Segmentation of two synthetic textures: From top to bottom and from left to right: Original image, averaged information scales, smoothed gradient of the scaled curvature, the outcome of the segmentation process after quantization into 4 levels.

In Figure 9 from top to bottom we see the original image, averaged information scales as depicted in the second step of the algorithm described above and the outcome of the segmentation process after smoothed by $3X3$ window and quantized to 7 levels. Notice the sensitivity of segmentation to texture.

Figure 9. Segmentation of mandrill image. From left to tight: Original image, averaged information scales, the outcome of the segmentation process after smoothed by $3X3$ window and quantized to 7 levels.

Figure 10 shows similar phenomena on an image of fabric with several textures. The figure shows the original image and the segmented outcome of the process. We can see good separation between the different textures.

The efficiency of the proposed the segmentation algorithm is highlighted on what might be called the ”semi -synthetic” (due to the regularity and quasi-periodicity of this natural image) of

the stairs – see Figure 11.

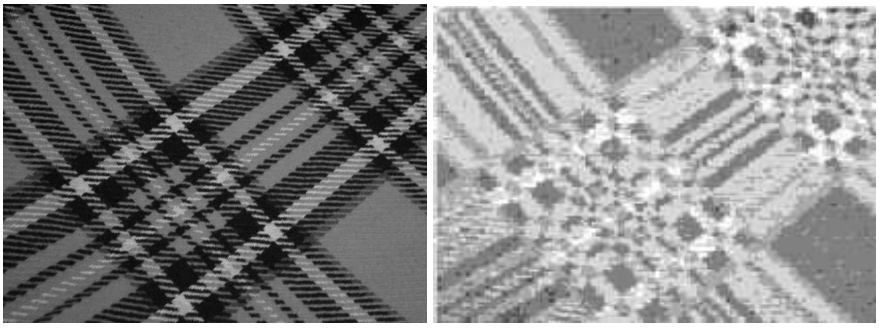


Figure 10. Segmentation of fabric image, window size and quantization level as in Figure 9. Although filter size is small one can easily see good differentiation between different structures along the fabric.

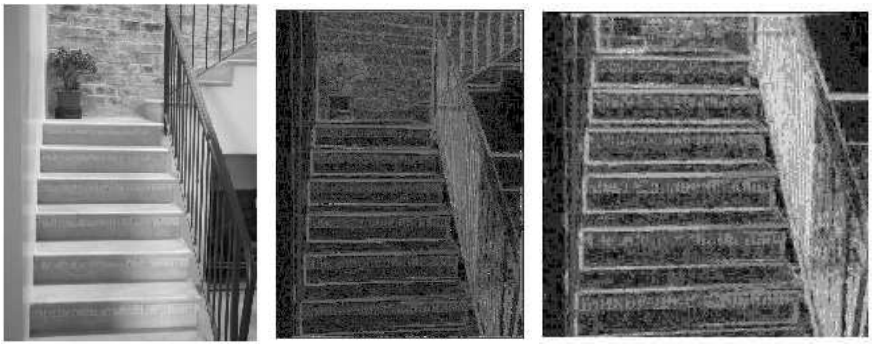


Figure 11. Segmentation of the stairs image: Original image (left), curvature computed using 3*X*3 windows (middle), detail of the segmented image using the same window size (right).

Finally, in Figure 12 the zebras are distinguished from the background of the image after their texture is segmented and separated from the texture of the background.

We conclude this Section with some preliminary comparison results. The briefness of this part is a direct consequence of the main goal of this paper, as stated in the introduction, and which we reiterate here briefly: Our essential objective is to continue and further develop the theoretical framework proposed in (Saucan *et al.*, 2010), that, in our opinion, allows, perhaps for the first time, to integrate, in a unique setting, the two common paradigms of Image Processing, namely the Harmonic Analysis/Wavelets and the Geometric (Graphics related) ones. Therefore, the present endeavor should be viewed rather as a feasibility check, rather that a bourn per se.

Nevertheless, some first comparisons were made, and we gaged our method by likening it with an established method for texture segmentation (Brox & Weickert, 2006) in conjunction with the use of the classical Gabor wavelets. Some of these results are presented in Figure 13.

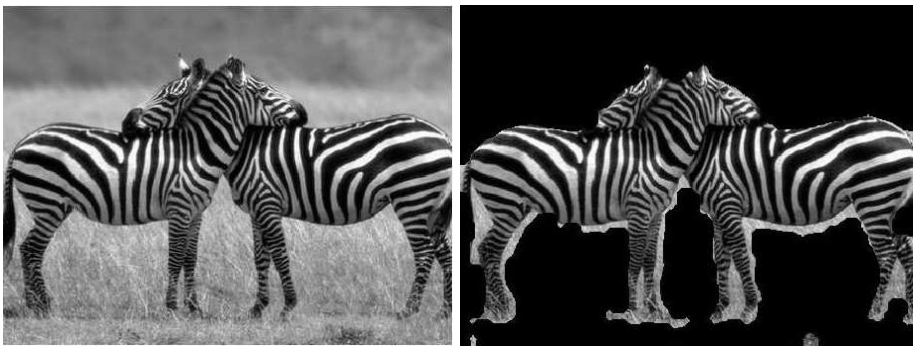


Figure 12. The zebras are extracted from the background of the original image. Extraction is based on the proposed texture segmentation.

5. Summary

The concept of scale is important for several image processing tasks. The calculation of scale on real life data usually relies on the convolution of the data with multi scale filter, where Gaussian derivatives are widespread. In this work we explored the relation between the concept of scale, to curvature. We have used the discrete definition proposed by Haantjes, and have established the theoretical exponential relation that is expected from geometry.

In addition, we propose to use the simpler curvature calculation as a means for automatic scale selection. In that respect, we show two interesting useful applications of the concept of curvature: as edge detectors and for the task of texture segmentation.

As we have noted in the introduction, scale and curvature are simply two manifests of the same physical phenomenon, it should evidently be that scale and curvature calculation can be inter-changed to accomplish the same tasks. However, while for scale only practical, intuitive, but not fully formalized definitions are given even in the most classical textbooks and other such authoritative sources, curvature – even metric one – is a classical, fully established and technical mathematical notion. We propose, therefore, in view of the remark above, to formally define scale by means of the Finsler-Haantjes curvature, at least in the purely theoretical setting. This is more relevant in the context of 2-dimensional (as well, of course, as higher dimensional ones), nonseparable signals, where a proper notion of scale is far less intuitive then in the 1-dimensional, classical, case.

Moreover, we suggest, to use this idea, not only for texture segmentation, but also in many other applications that make use of scale analysis of signals in general and of images in particular. Due to the efficiency of the computation of Haantjes curvature relative to, for instance, wavelets and Gabor functions computations in many scales, we can regard for applications both in image processing, as well as machine vision, where usually efficiency is essential. Just to name a few, we suggest the following,

a Compression and compress sensing.

b Detection of key points in images and registration.

c Scale space representation.

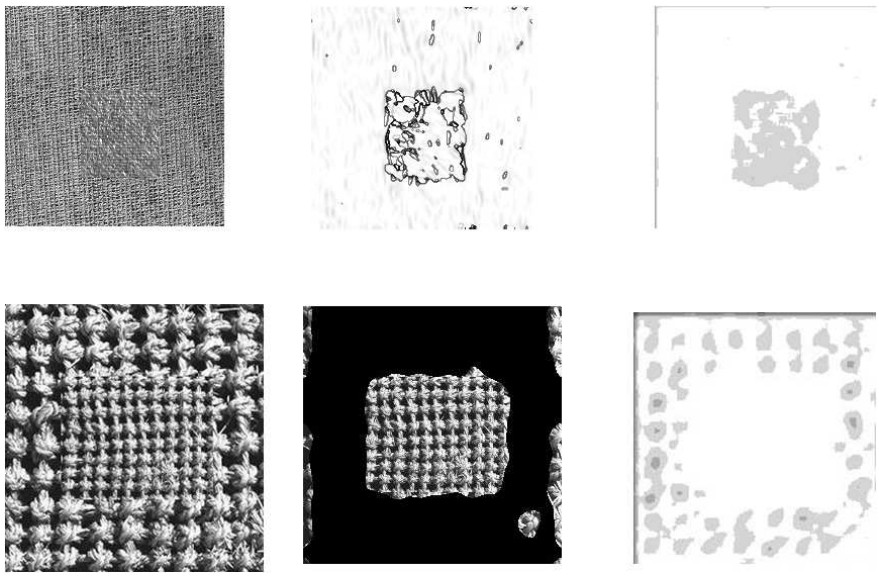


Figure 13. Texture segmentation: Gabor wavelets based segmentation (above) versus curvature based segmentation (below).

d Adaptive edge detection.

e Object recognition.

As for further possible research issues we believe that what we have presented in this paper is actually the tip of an iceberg as far the scale-curvature connection is concerned. Again, to name a few we can mention the following directions,

a Automatic scale selection in the sense of pointing out automatically a scale up to which one can apply analysis-synthesis process with a guaranteed accuracy

b Use additional information which is obtained during the process for additional tasks. As the curvature is computed we gain information about all scales at which, the curvature jumps above a predefined threshold. The method presented herein only makes use of the average of all these scales however one can employ this information for an adaptive scale selection making use of the relevant information of each of these.

c In addition to the above, there is also information about the various directions at which curvature is computed, which is obtained during the process and this information can certainly be exploited for a variety of implementations such as those mentioned above.

d Can we use the curvature-scale relation in order to gain information about the "adequacy" of a certain wavelet family to a given signal? For instance, it is often asked, is it beneficial, in any way, to decompose say, a natural image using a specific wavelets family over the others, say, for obtaining sparse representation? We hope that some answers can be given for this challenging question, via the scale-curvature analysis. We suggest to account for e.g., the exponential decay

of curvature as a function of scale, for the given signal, and then looking for the wavelet family with most similar behavior. In fact this particular question was the one that motivated this line of research.

Acknowledgments

Emil Saucan's research was partly supported by Israel Science Foundation Grants 221/07 and 93/11. and by European Research Council under the European Communitys Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement nº [203134].

References

Antoine, J.P. and L. Jaques (2003). Measuring a curvature radius with directional wavletes. In: *GROUP 24: Physical and Mathematical Aspects of Symmetries, Inst. Phys. Conf. Series*. In J-P. Gazeau, R. Kerner, J-P. Antoine, S. Metens, J-Y. Thibon, (Eds.) **173**(8), 899–904.

Bay, H., T. Tuytelaars and L. Van Gool (2006). SURF: Speeded Up Robust Features. *Lecture Notes in Computer Science* **3951**, 404–417.

Beauchamp, K.G. (1975). *Walsh Functions and Their Applications*. London Academic Press.

Blumenthal, L.M. and K. Menger (1970). *Studies in Geometry*. Freeman & co., San Francisco.

Brox, T. and J. Weickert (2006). A TV flow based local scale estimate and its application to texture discrimination. *Journal of Visual Communication and Image Representation*. **17**(5), 1053–1073.

Cohen, L. (1993). The Scale Representation. *IEEE Trans. Signal Processing*. **41**(12), 3275–3292.

do Carmo, M.P (1976). *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Englewood Cliffs, NJ.

Florack (1997). Image Structure. *Computational Imaging and Vision* **10**. Kluwer Academic Publishers, Dordrecht.

Florack, L., B.M. ter Haar Romeny, J.J. Koenderink and M.A. Viergever (1992). Scale and the differential structure of images. *Image Vision Comput.* **10**(6), 376–388.

Koenderink, J.J. (1984). The structure of images. *Biological Cybernetics*. **50**, 363–370.

Lindeberg, T. (1998). Feature detection with automatic scale selection. *International Journal of Computer Vision*. **30**, 77–116.

Lounsbery, J.M., A.D. DeRose and J. Warren (1997). Multiresolution Analysis For Surfaces Of Arbitrary Topological Type. *ACM Transactions on Graphics*. **16**(1), 34–73.

Lowe, D.G. (1999). Object recognition from local scale-invariant features. *ICCV '99*. **2**, 1150–1157.

Meyer, Y. (1993). *Wavelets : Algorithms & Applications*. SIAM, University of Michigan, MI.

Petersen, P. (1998). *Riemannian Geometry*. Springer-Verlag, New York.

Saucan, E. (2006). Curvature – Smooth, Piecewise-Linear and Metric. In: *What is Geometry?, Advanced Studies in Mathematics and Logic*. G. Sica, (Ed.), pp. 237-268

Saucan, E. and E. Appleboim (2005). Curvature Based Clustering for DNA Microarray Data Analysis. *Lecture Notes in Computer Science* **3523**, 405–412.

Saucan, E. and E. Appleboim (2009). Metric Methods in Surface Triangulation. *Lecture Notes in Computer Science* **5654**, 335–355.

Saucan, E., C. Sagiv and E. Appleboim (2010). Geometric Wavelets for Image Processing: Metric Curvature of Wavelets. In: *Proceedings of SampTA 2009, France, Marseille, May 18-22, 2009*. pp 85–89.

Strömberg, J.O. (1983). A modified Franklin system and high order spline systems on $\mathbb{R}^n$ as unconditional bases for Hardy spaces. In: *Conference on Harmonic Analysis in honor of A. Zygmund*. Wadeworth International Group, Belmont, CA. W. Beckner, (Ed.), pp 475-494.

Valette, S. and R. Prost (2004). Wavelet-Based Multiresolution Analysis Of Irregular Surface Meshes. *IEEE Transaction on Visualization and Computer Graphics*. **10**(2), 113–122.

Witkin, A.P. (1983). Scale-space filtering. In: *Proceedings of the Eighth international joint conference on Artificial intelligence*. Vol. 2, pp. 1019–1022.

Witkin, A.P. (1983). Scale Space Filtering: A New Approach to Multi-Scale Descriptions. In: *Proc. 8th Int. Joint Conf. Art. Intell., Germany, Karlsruhe*. pp. 1019–1022.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 56–64

Third Order Boundary Value Problem with Integral Condition at Resonance

Assia Guezane-Lakoud[a,*], Assia Frioui[b]

[a]*Laboratory of Advanced Materials, Faculty of Sciences Badji Mokhtar-Annaba University P. O. Box 12, 23000 Annaba, Algeria*
[b]*Laboratory of applied mathematics and modeling, Department of Mathematics, University 08 Mai45-Guelma, P. O. Box 401, Guelma 24000, Algeria*

Abstract

This paper deals with a class of third order boundary value problem with integral condition at resonance. Some existence results are obtained by using the coincidence degree theory of Mawhin.

Keywords: Fixed point theorem, coincidence degree theory of Mawhin, third order boundary value problem, integral condition, Fredholm operators, resonance.
2010 MSC: 34B15, 34B18, 34G20.

1. Introduction

Let us consider the following third-order differential equation:

$$x'''(t) = f\left(t, x(t), x'(t)\right), 0 < t < 1, \tag{1.1}$$

subject to the following nonlocal conditions

$$x(0) = x''(0) = 0, x(1) = \frac{2}{\eta^2}\int_0^\eta x(t)\,dt, \eta \in (0,1), \tag{1.2}$$

where $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ is a Carathéodory function, and $\eta \in (0,1)$. We say that the boundary value problem (1.1), (1.2) is a resonance problem if the linear equation $Lx = x'''$, with the boundary value conditions (1.2) has a non-trivial solution i.e., $\dim \ker L \geq 1$.

*Corresponding author
Email addresses: a_guezane@yahoo.fr (Assia Guezane-Lakoud), frioui.assia@yahoo.fr (Assia Frioui)

The research of ordinary differential equations with nonlocal conditions plays a very important role in both theory and applications. It is widely used in describing a large number of physical, biological and chemical phenomena. Moreover, the theory of boundary-value problems with integral boundary conditions arises in different areas of applied mathematics and physics. For example, heat conduction, chemical engineering, underground water flow, thermo-elasticity, and plasma physics can be reduced to the nonlocal problems with integral boundary conditions. In recent years, the multi-point boundary value problems at resonance for second order, third order ordinary differential equations have been extensively studied and many excellent results have been obtained, for instance, see (Feng & Webb, 1997*a*), (Feng & Webb, 1997*b*), (Gupta, 1995), (Gupta & Tsamatos, 1994), (Liu & Yu, 2002), (Liu, 2003), (Liu & Zhao, 2007), (Kosmatov, 2006), (Du, 2008), (Du & Ge, 2005), (Ma, 2005), (Nagle & Pothoven, 1995), (Xue & Ge, 2004). (see, also, (Y. Liu, 2005), (X. Lin, 2009), (H Zhang & Chen, 2009)). However, to our knowledge, the corresponding results for third-order with integral boundary conditions, are rarely seen (see, for example, (X. Lin & Meng, 2011), (Karakostas & Tsamatos, 2002), (Yang, 2006); (A. Yang, 2011) and references therein). (Meng & Du, 2010) studied the following second-order multi-point boundary value problem at resonance:

$$\begin{cases} x''(t) = f(t, x(t), x'(t)) + e(t), t \in (0,1), \\ x(0) = \sum_{i=1}^{m} \alpha_i x(\xi_i), \ x'(1) = \sum_{j=1}^{n} \beta_j x\left(\eta_j\right), \end{cases}$$

where $f : [0,1] \times \mathbb{R}^2 \to \mathbb{R}$ is a Carathéodory function, $e \in L^1[0,1], 0 < \xi_1 < .... < \xi_m < 1$, $\alpha_i \in \mathbb{R}$, $i = 1, 2, ...m, m \geq 2$ and $0 < \eta_1 < ..... < \eta_n < 1, \beta_j \in \mathbb{R}$, $j = 1, ...n, n \geq 1$. By using coincidence degree of Mawhin the authors obtain many excellent results about the existence of solutions for the above problem under the resonance conditions $\sum_{i=1}^{m} \alpha_i = \sum_{j=1}^{n} \beta_j = 1$ and $\sum_{i=1}^{m} \alpha_i \xi_i = 0$.

By using coincidence degree of Mawhin (Lin & Meng, 2011), established the existence of solutions for the following third-order multi-point boundary value problem at resonance

$$\begin{cases} x'''(t) = f(t, x(t), x'(t), x''(t)), 0 < t < 1 \\ x''(0) = \sum_{i=1}^{m} \alpha_i x''(\xi_i), \quad x'(0) = 0, x(1) = \sum_{j=1}^{n} \beta_j x\left(\eta_j\right), \end{cases}$$

where $0 < \xi_1 < .... < \xi_m < 1$, $\alpha_i \in \mathbb{R}$, $i = 1, ...m, m \geq 1$ and $0 < \eta_1 < ..... < \eta_n < 1, \beta_j \in \mathbb{R}$, $j = 1, ...n, n \geq 2$, and $f : [0,1] \times \mathbb{R}^3 \to \mathbb{R}$ is a continuous function.

More recently, (X. Zhang & Ge, 2009) studied the following nonlocal boundary value problem:

$$\begin{cases} x''(t) = f(t, x(t), x'(t)) + e(t), t \in (0,1) \\ x'(0) = \int_0^1 h(t) x'(t)\, dt, x'(1) = \int_0^1 g(t) x'(t)\, dt, \end{cases}$$

where $f, g \in C([0,1], [0,\infty))$. Especially by using the coincidence degree of Mawhin, and under the resonance conditions $\int_0^1 h(t)\, dt = 1$, and $\int_0^1 g(t)\, dt = 1$, the authors proved at least one solution.of the boundary value problem.

The purpose of this paper is to study the existence of solutions for nonlocal boundary value problem $(1,1), (1,2)$ at resonance and establish an existence theorem . Our method is based upon the coincidence degree theory of (Mawhin, 1979).

2. Main results

We first recall some notation and an abstract existence result (Mawhin, 1979).

Let X, Y be two real Banach spaces and let $L : domL \subset X \to Y$ be a linear operator which is Fredholm map of index zero and $P : X \to X$, $Q : Y \to Y$ be continuous projectors such that $ImP = KerL$, $KerQ = ImL$ and $X = KerL \oplus KerP$, $Y = ImL \oplus ImQ$. It follows that $L|_{domL \cap KerP} : domL \cap KerP \to ImL$ is invertible, we denote the inverse of that map by K_P. Let Ω be an open bounded subset of X such that $domL \cap \Omega \neq \emptyset$, the map $N : X \to Y$ is said to be L-compact on $\overline{\Omega}$ if the map $QN\left(\overline{\Omega}\right)$ is bounded and $K_P(I - QN) : \overline{\Omega} \to X$ is compact. To obtain our existence results we use the following fixed point theorem of (Mawhin, 1979).

Theorem 2.1. *Let be L a Fredholm operator of index zero and N be L− compact on* $\overline{\Omega}$. *Assume that the following conditions are satisfied:*

i) $Lx \neq \lambda Nx$ *for every* $(x, \lambda) \in [(domL \backslash KerL) \cap \partial\Omega] \times (0, 1)$.

ii) $Nx \notin \text{Im } L$ *for every* $x \in KerL \cap \partial\Omega$.

iii) $\deg(QN|_{KerL}, \Omega \cap KerL, 0) \neq 0$, *where* $Q : Y \to Y$ *is a projection as above with* $\text{Im } L = KerQ$.

Then the equation $Lx = Nx$ *has at least one solution in* $domL \cap \overline{\Omega}$.

In the following, we shall use the classical spaces $C[0,1]$, $C^1[0,1]$, $C^2[0,1]$ and $L^1[0,1]$. For $x \in C^2[0,1]$, we use the norm $\|x\| = \max\{\|x\|_\infty, \|x'\|_\infty\}$ where $\|x\|_\infty = \max_{t\in[0,1]} |x(t)|$ and denote the norm in $L^1[0,1]$ by $\|\cdot\|_1$. We will use the Sobolev space $W^{3,1}(0,1)$ which is defined by $W^{3,1}(0,1) = \{x : [0,1] \to \mathbb{R} : x, x', x'' \text{ are absolutely continuous on } [0,1] \text{ with } x''' \in L^1[0,1]\}$.

Let $X = C^2[0,1]$, $Y = L^1[0,1]$, L is the linear operator from $domL \subset X$ to Y with $domL = \left\{x \in W^{3,1}(0,1) : x(0) = x''(0) = 0,\ x(1) = \frac{2}{\eta^2}\int_0^\eta x(t)\,dt\right\}$ and $Lx = x'''$, $x \in \text{dom } L$. We define $N : X \to Y$ by setting

$$Nx = f(t, x(t), x'(t)), t \in (0,1).$$

Then the BVP (1.1) and (1.2) can be written as $Lx = Nx$.

Theorem 2.2. *Assume that the following conditions are satisfied:*

1) There exists functions $\alpha, \beta, \gamma \in L^1[0,1]$, *such that for all* $(x, y) \in \mathbb{R}^2$, $t \in [0,1]$ *then*

$$|f(t,x,y)| \leq \alpha(t)|x| + \beta(t)|y| + \gamma(t). \tag{2.1}$$

2) There exists a constant $M > 0$, *such that for* $x \in domL$, *if* $|x'(t)| > M$ *for all* $t \in [0,1]$, *then*

$$\int_0^1 (1-s)^2 f(s, x(s), x'(s))\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 f(s, x(s), x'(s))\,ds \neq 0. \tag{2.2}$$

3) There exists a constant $M^* > 0$, *such that for any* $x(t) = bt \in KerL$ *with* $|b| > M^*$, *either*

$$b\left[\int_0^1 (1-s)^2 f(s, b(s), b)\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 f(s, b(s), b)\,ds\right] < 0, \tag{2.3}$$

or else

$$b\left[\int_0^1 (1-s)^2 f(s, b(s), b)\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 f(s, b(s), b)\,ds\right] > 0. \tag{2.4}$$

then BVP $(1,1)$ *and* $(1,2)$ *has at least one solution in* $C^2[0,1]$, *provided*

$$\|\alpha\| + \|\beta\| < \frac{1}{2}. \tag{2.5}$$

2.1. Proof of Theorem 2.2

For the proof of Theorem 2.2 we shall apply Theorem 2.1 and the following lemmas.

Lemma 2.1. *The operator $L : domL \subset X \to Y$ is a Fredholm operator of index zero. Furthermore, the linear projector operator $Q : Y \to Y$ can be definied by*

$$Qy(t) = k\left[\int_0^1 (1-s)^2 y(s)\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 y(s)\,ds\right]t,$$

where $k = 60/5 - 2\eta^3$ and the linear operator $K_P : \text{Im } L \to dom\, L \cap Ker\, P$ can be written by

$$K_p y(t) = \frac{1}{2}\int_0^t (t-s)^2 y(s)\,ds, \forall y \in \text{Im } L.$$

Furthermore

$$\left\|K_p y\right\| \le \|y\|_1, \forall y \in \text{Im } L.$$

Proof. It is clear that

$$kerL = \{x \in \text{dom } L : x = bt,\ b \in \mathbb{R},\ t \in [0,1]\} \simeq \mathbb{R}.$$

Now we show that

$$\text{Im } L = \left\{ y \in Y : \int_0^1 (1-s)^2 y(s)\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 y(s)\,ds = 0 \right\}. \tag{2.6}$$

The problem

$$x''' = y \tag{2.7}$$

has a solution $x(t)$ that satisfies the conditions $x(0) = x''(0) = 0,\ x(1) = \frac{2}{\eta^2}\int_0^\eta x(t)\,dt$, if and only if

$$\int_0^1 (1-s)^2 y(s)\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 y(s)\,ds = 0. \tag{2.8}$$

In fact from (2.7) we have

$$x(t) = x''(0)\frac{t^2}{2} + x'(0)t + x(0) + \frac{1}{2}\int_0^t (t-s)^2 y(s)\,ds = x'(0)t + \frac{1}{2}\int_0^t (t-s)^2 y(s)\,ds.$$

According to $x(1) = \frac{2}{\eta^2}\int_0^\eta x(t)\,dt$,we obtain

$$\int_0^1 (1-s)^2 y(s)\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 y(s)\,ds = 0.$$

On the other hand, if (2.8) holds, setting

$$x(t) = bt + \frac{1}{2}\int_0^t (t-s)^2 y(s)\,ds,$$

where b is an arbitrary constant, then $x(t)$ is a solution of (2.7). Hence (2.6) holds.
Setting

$$Ry = \int_0^1 (1-s)^2 y(s)\,ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 y(s)\,ds,$$

define $Qy(t) = k.(Ry).t$, it is clear that $\dim \mathrm{Im}Q = 1$. We have

$$Q^2 y = Q(Qy) = k(k.Ry)\left(\int_0^1 (1-s)^2 sds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 sds\right)t = (kRy)\,t = Qy,$$

which implies that the operator Q is a projector. Furthermore, $\mathrm{Im}\,L = kerQ$.
Let $y = (y - Qy) + Qy$, where $y - Qy \in KerQ = \mathrm{Im}\,L$, $Qy \in \mathrm{Im}Q$. It follows from $KerQ = \mathrm{Im}\,L$ and $Q^2 y = Qy$ that $\mathrm{Im}\,Q \cap \mathrm{Im}\,L = \{0\}$. Then, we have $Y = \mathrm{Im}\,L \oplus \mathrm{Im}\,Q$. Since $\dim KerL = 1 = \dim \mathrm{Im}\,Q = co\dim \mathrm{Im}\,L = 1$, L is a Fredholm map of index zero.
Now we define a projector P from X to X by setting

$$Px(t) = x'(0)\,t.$$

Then the generalized inverse $K_P : \mathrm{Im}\,L \to domL \cap KerP$ of L can be written by

$$K_p = \frac{1}{2}\int_0^t (t-s)^2 y(s)\,ds.$$

Obviously, $\mathrm{Im}\,P = KerL$ and $P^2 x = Px$. It follows from $x = (x - Px) + Px$ that $X = KerP + KerL$. By simple calculation, we can get that $KerL \cap KerP = \{0\}$. Then $X = KerL \oplus KerP$. From the definitions of P and K_P it is easy to see that the generalized inverse of L is K_P. In fact, for $y \in \mathrm{Im}\,L$, we have

$$\left(LK_p\right)y(t) = \left[\left(K_p y\right)t\right]''' = y(t),$$

and for $x \in domL \cap kerP$, we know

$$\left(K_p L\right)x(t) = \left(K_p\right)x'''(t) = \frac{1}{2}\int_0^t (t-s)^2 x'''(s)\,ds = x(t) - x(0) - x'(0)\,t - \frac{1}{2}x''(0)\,t^2,$$

in view of $x \in domL \cap kerP$, $x(0) = x''(0) = 0$ and $Px = 0$, thus

$$\left(K_p L\right)x(t) = x(t).$$

This shows that $K_p = \left(L|_{\mathrm{dom}L \cap \mathrm{ker}P}\right)^{-1}$. Also we have

$$\left\|K_p y\right\|_\infty \le \int_0^1 (1-s)^2 |y(s)|\,ds \le \int_0^1 |y(s)|\,ds = \|y\|_1,$$

and from $\left(K_p y\right)'(t) = \int_0^1 (1-s)\, y(s)\, ds$, we obtain

$$\left\|\left(K_p y\right)'\right\|_\infty \leq \int_0^1 (1-s)\,|y(s)|\, ds \leq \int_0^1 |y(s)|\, ds = \|y\|_1$$

then $\left\|K_p y\right\| \leq \|y\|_1$. This completes the proof of Lemma 2.1. □

Lemma 2.2. *Let* $\Omega_1 = \{x \in domL \backslash KerL : Lx = \lambda Nx, for\ some\ \lambda \in [0,1]\}$. *Then* Ω_1 *is bounded.*

Proof. Suppose that $x \in \Omega_1$, and $Lx = \lambda Nx$. Thus $\lambda \neq 0$ and $QNx = 0$, so it yields

$$\int_0^1 (1-s)^2 f(s, x(s), x'(s))\, ds - \frac{2}{3\eta^2} \int_0^\eta (\eta - s)^3 f(s, x(s), x'(s))\, ds = 0.$$

Thus, by condition (2), there exists $t_0 \in [0,1]$, such that $|x'(t)| \leq M$. In view of

$$x'(0) = x'(t_0) - \int_0^{t_0} x''(t)\, dt, x''(t) = x''(0) + \int_0^t x'''(s)\, ds,$$

then, we have

$$|x'(0)| \leq M + \int_0^1 \left(\int_0^1 |x'''(s)\, ds| \right) dt = M + \|x'''\|_1 = M + \|Lx\|_1 \leq M + \|Nx\|_1 . \tag{2.9}$$

Again for $x \in \Omega_1$, $x \in domL \backslash KerL$, then $(I-P)x \in domL \cap KerP$ and $LPx = 0$, thus from Lemma 3, we know

$$\|(I-P)x\| = \left\|K_p L(I-Px)\right\| \leq \|L(I-Px)\|_1 = \|Lx\|_1 \leq \|Nx\|_1 . \tag{2.10}$$

From (2.9) and (2.10), we have

$$\|x\| \leq \|Px\| + \|(I-P)x\| = |x'(0)| + \|(I-P)x\| \leq M + 2\|Nx\|_1 . \tag{2.11}$$

From (2.1) and (2.11), we obtain

$$\|x\| \leq 2\left[\|\alpha\|_1 \|x\|_\infty + \|\beta\|_1 \|x'\|_\infty + \|\gamma\|_1 + \frac{M}{2}\right]. \tag{2.12}$$

Thus, from $\|x\|_\infty \leq \|x\|$ and (2.12) we have

$$\|x\|_\infty \leq \frac{2}{1 - 2\|\alpha\|_1}\left[\|\beta\|_1 \|x'\|_\infty + \|\gamma\|_1 + \frac{M}{2}\right]. \tag{2.13}$$

From $\|x'\|_\infty \leq \|x\|$, and (2.12) and (2.13), one has

$$\|x'\|_\infty \left[1 - \frac{2\|\beta\|_1}{1 - 2\|\alpha\|_1}\right] \leq \frac{2}{1 - 2\|\alpha\|_1}\left[\|\gamma\|_1 + \frac{M}{2}\right].$$

Therefore,

$$\|x'\|_\infty \left[\frac{1-2\|\alpha\|_1-2\|\beta\|_1}{1-2\|\alpha\|_1}\right] \leq \frac{1}{1-2\|\alpha\|_1}\left[2\|\gamma\|_1+M\right].$$

i.e.,

$$\|x'\|_\infty \leq \frac{2\left[\|\gamma\|_1+\frac{M}{2}\right]}{1-2\|\alpha\|_1-2\|\beta\|_1} = M_1. \tag{2.14}$$

From (2.14), there exists $M_1 > 0$, such that

$$\|x'\|_\infty \leq M_1, \tag{2.15}$$

thus from (2.15) and (2.13), there exists $M_2 > 0$, such that

$$\|x\|_\infty \leq M_2. \tag{2.16}$$

Hence

$$\|x\| = \max\{\|x\|_\infty, \|x'\|_\infty\} \leq \max\{M_1, M_2\}.$$

Again from (2.1), (2.15) and (2.16), we have

$$\|x'''\|_1 = \|Lx\|_1 \leq \|Nx\|_1 \leq \|\alpha\|_1 M_2 + \|\beta\|_1 M_1 + \|\gamma\|_1.$$

So Ω_1 is bounded. □

Lemma 2.3. *The set* $\Omega_2 = \{x \in KerL : Nx \in \text{Im } L\}$ *is bounded.*

Proof. Let $x \in \Omega_2$, then $x \in KerL = \{x \in domL : x = bt,\ b \in \mathbb{R},\ t \in [0,1]\}$, and $QNx = 0$, therefore

$$\int_0^1 (1-s)^2 f(s, bs, b)\, ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 f(s, bs, b)\, ds = 0.$$

From condition (2) of Theorem 2.2, $\|x\|_\infty = |b| \leq M$, so $\|x\| = |b| \leq M$, thus Ω_2 is bounded. □

Lemma 2.4. *If the first part of condition (3) of Theorem 2.2 holds, then*

$$b\frac{60}{5-2\eta^3}\left[\int_0^1 (1-s)^2 f(s, b(s), b)\, ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 f(s, b(s), b)\, ds\right] < 0, \tag{2.17}$$

for all $|b| > M^*$. *Let* $\Omega_3 = \{x \in KerL : -\lambda Jx + (1-\lambda)QNx = 0,\ \ \lambda \in [0,1]\ \}$ *where* $J : KerL \to$ Im Q *is the linear isomorphism given by* $J(bt) = bt$, $\forall b \in \mathbb{R}$, $t \in [0,1]$. *Then* Ω_3 *is bounded.*

Proof. Suppose that $x = b_0 t \in \Omega_3$, then we obtain

$$\lambda b_0 = (1-\lambda)\frac{60}{5-2\eta^3} \times \left(\int_0^1 (1-s)^2 f(s, b(s), b)\, ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 f(s, b(s), b)\, ds\right).$$

If $\lambda = 1$, then $b_0 = 0$. Otherwise, if $|b_0| > M^*$, then in view of (2.17) one has $\lambda b_0^2 = b_0(1-\lambda)\frac{60}{5-2\eta^3} \times \left(\int_0^1 (1-s)^2 f(s, b(s), b)\, ds - \frac{2}{3\eta^2}\int_0^\eta (\eta - s)^3 f(s, b(s), b)\, ds\right) < 0,$

which contradicts the fact that $\lambda b_0^2 \geq 0$. Then $|x| = |b_0 t| \leq |b_0| \leq M^*$, we obtain $\|x\| \leq M^*$, therefore $\Omega_3 \subset \{x \in KerL : \|x\| \leq M^*\}$ is bounded.

If $\lambda = 0$, it yields

$$\int_0^1 (1-s)^2 f(s, b(s), b)\, ds - \frac{2}{3\eta^2} \int_0^\eta (\eta - s)^3 f(s, b(s), b)\, ds = 0,$$

taking condition (2) of Theorem 2.2 into account, we obtain $\|x\| = |b| \leq M^*$. □

Lemma 2.5. *If the second part of condition (3) of Theorem 2.2 holds, then*

$$b \frac{60}{5 - 2\eta^3} \left[\int_0^1 (1-s)^2 f(s, b(s), b)\, ds - \frac{2}{3\eta^2} \int_0^\eta (\eta - s)^3 f(s, b(s), b)\, ds \right] > 0, \tag{2.18}$$

for all $|b| > M^*$. Let $\Omega_3 = \{x \in KerL : \lambda Jx + (1-\lambda) QNx = 0, \ \lambda \in [0,1]\}$, here J is defined as in Lemma 2.4. Similar to the above argument, we can verify that Ω_3 is bounded.

Now the proof of Theorem 2.2 is a consequence of Theorem 2.1 and the above lemmas.

Proof. **of Theorem 2.2**. Let Ω to be an open bounded subset of X such that $\cup_{i=1}^3 \overline{\Omega}_i \subset \Omega$. By using the Arzela- Ascoli theorem, we can prove that $K_P(I - QN) : \overline{\Omega} \to X$ is compact, thus N is L-compact on $\overline{\Omega}$. Then by Lemmas 2.2 and 2.3, we have

i) $Lx \neq \lambda Nx$ pour tout $(x, \lambda) \in [(domL \backslash KerL) \cap \partial\Omega] \times (0,1)$.

ii) $Nx \notin \text{Im } L$ pour tout $x \in KerL \cap \partial\Omega$.

iii) Let $H(x, \lambda) = \pm \lambda Jx + (1-\lambda) QNx = 0$.

According to Lemmas 2.4 and 2.5, we know that $H(x, \lambda) \neq 0$ for every $x \in KerL \cap \partial\Omega$. Thus, by the homotopy property of degree, $\deg(QN|_{KerL}, \Omega \cap KerL, 0) = \deg(H(\cdot, 0), \Omega \cap KerL, 0) = \deg(H(\cdot, 1), \Omega \cap KerL, 0) = \deg(\pm J, \Omega \cap KerL, 0) \neq 0$. Then by Theorem 2.1, $Lx = Nx$ has at least one solution in $domL \cap \overline{\Omega}$, so the BVP (1.1), (1, 2) has at least one solution in $C^2[0,1]$. The proof is complete. □

References

A. Yang, B. Sun, W. Ge (2011). Existence of positive solutions for self-adjiont boundary-value problems with integral boundary condition at resonance. *Electron. J. Differential Equations* **11**, 1–8.

Du, Z (2008). Solvability of functional differential equations with multi-point boundary value problem at resonance. *Comput. Math. Appl* **55**, 2653–2661.

Du, Z., X. Lin and W. Ge (2005). On a third-order multi-point boundary value problem at resonance. *J. Math. Anal. Appl* **302**, 217–229.

Feng, W. and J. R. L. Webb (1997*a*). Solvability of m-point boundary value problems with nonlinear growth. *J. Math. Anal. Appl* **212**, 467–480.

Feng, W. and J. R. L. Webb (1997*b*). Solvability of three-point boundary value problems at resonance. *Nonlinear Anal. Theory, Methods and Appl* **30**, 3227–3238.

Gupta, C.P (1995). A second order m-point boundary value problem at resonance. *Nonlinear Anal* **24**, 1036–1046.

Gupta, C.P., S.K. Ntouyas and P.Ch. Tsamatos (1994). On an m-point boundary-value problem for second-order ordinary differential equations. *Nonlinear Anal* **23**, 1427–1436.

H Zhang, W Liu, J Zhang and T Chen (2009). Existence of solutions for three- point boundary value problem. *J. Appl. Math. & Informatics* **27**(5-6), 35–51.

Karakostas, G. L. and P. Ch. Tsamatos (2002). Sufficient conditions for the existence of nonnegative solutions of a nonlocal boundary value problem. *Appl. Math. Letters* **15**(4), 401–407.

Kosmatov, N (2006). A multi-point boundary value problem with two critical conditions. *Nonlinear Analysis: Theory, Methods & Applications* **65**, 622–633.

Lin, X., Z. Du. and F. Meng (2011). A note on a third-order multi-point boundary value problem at resonance. *Math. Nachr.* **284**(13), 1690 – 1700.

Liu, B (2003). Solvability of multi-point boundary value problem at resonance (ii). *Appl. Math. Comput* **136**, 353–377.

Liu, B. and J. S. Yu (2002). Solvability of multi-point boundary value problems at resonance (i). *Indian J. Pure Appl. Math* **34**, 475–494.

Liu, B. and Z. Zhao (2007). A note on multi-point boundary value problems. *Nonlinear Anal* **67**, 2680–2689.

Ma, R (2005). Multiplicity results for a third order boundary value problem at resonance, nonlinear anal. *Nonlinear Anal* **32**, 493–499.

Mawhin, J. (1979). *Topological degree methods in nonlinear boundary value problems*. NSFCBMS Regional Conference Series in Mathematics, American Mathematical Society, Providence, RI.

Meng, F. and Z. Du (2010). Solvability of a second order multi-point boundary value problem at resonance. *Appl. Math. Comput* **208**, 23–30.

Nagle, R. K. and K. L. Pothoven (1995). On a third-order nonlinear boundary value problems at resonance. *J. Math. Anal. Appl* **195**, 148–159.

X. Lin, W. Liu (2009). A nonlinear third-order multi-point boundary value prob- lems in the resonance case. *J. App Math comput* **29**, 35–51.

X. Lin, Z. Du and F. Meng (2011). Existence of solutions to a nonlocal boundary value problem with nonlinear growth. *Boundary Value Problems* **2011**, 15pp.

X. Zhang, M. Feng and W. Ge (2009). Existence result of second-order differential equations with integral boundary conditions at resonance. *J. Math. Anal. Appl* **353**(1), 311–319.

Xue, C., Z. Du. and W. Ge (2004). Solutions to m-point boundary value problems of third order ordinary differential equations at resonance. *J. Appl. Math. Comput* **17**(1-2), 299–244.

Y. Liu, W Ge (2005). Solution of multi-point boundary value problems for higher-order differential equations at resonance(iii). *Tamkang Journal of Mathematics* **36**(2), 119–130.

Yang, Z. (2006). Positive solutions of a second-order integral boundary-value problem. *J. Math. Anal. Appl* **321**, 751–765.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 65–84

The $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-Univexities of Higher-Orders with Applications to Parametric Duality Models in Minimax Fractional Programming

Ram U Verma[a,*]

[a]*International Publications USA, 3400 S Brahma Blvd Suite 31B, Kingsville, TX 78363, USA*

Abstract

Based on the recently introduced (see (Verma, 2012)) major higher order generalizations ($\mathcal{G}$, β, ϕ, $h(\cdot,\cdot)$, ρ, θ) - univexities, several second-order parametric duality models for a semiinfinite minimax fractional programming problem are developed with appropriate duality results under various generalized second-order ($\mathcal{G}$, β, ϕ, $h(\cdot,\cdot)$, ρ, θ) - univexity assumptions. The obtained results encompass a large variety of investigations on generalized univexities and their extensions in the literature.

Keywords: Semiinfinite programming, minimax fractional programming, generalized second-order univex functions, infinitely many equality and inequality constraints, dual problems, duality theorems.
2010 MSC: 49N15, 90C26, 90C30, 90C32, 90C45, 90C47.

1. Introduction

In this paper, we intend to establish some results on second-order duality under various generalized $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-univexity assumptions for the semiinfinite discrete minimax fractional programming problem of the form:

$$(P) \qquad \text{Minimize} \ \max_{1\le i\le p} \frac{f_i(x)}{g_i(x)}$$

subject to

$$G_j(x,t) \le 0 \ \text{ for all } t \in T_j, \quad j \in \underline{q} = \{1,2,\cdots,q\},$$
$$H_k(x,s) = 0 \ \text{ for all } s \in S_k, \quad k \in \underline{r} = \{1,2,\cdots,r\},$$
$$x \in X,$$

*Corresponding author
Email address: verma99@msn.com (Ram U Verma)

where p, q, and r are positive integers, X is a nonempty open convex subset of $\mathbb{R}^n$ (n-dimensional Euclidean space), for each $j \in \underline{q} = \{1, 2, \ldots, q\}$ and $k \in \underline{r} = \{1, 2, \ldots, r\}$, T_j and S_k are compact subsets of complete metric spaces, for each $i \in \underline{p}$, f_i and g_i are twice continuously differentiable real-valued functions defined on X, for each $j \in \underline{q}$, $z \to G_j(z, t)$ is a twice continuously differentiable real-valued function defined on X for all $t \in T_j$, for each $k \in \underline{r}$, $z \to H_k(z, s)$ is a twice continuously differentiable real-valued function defined on X for all $s \in S_k$, for each $j \in \underline{q}$ and $k \in \underline{r}$, $t \to G_j(x, t)$ and $s \to H_k(x, s)$ are continuous real-valued functions defined, respectively, on T_j and S_k for all $x \in X$, and for each $i \in \underline{p}$, $g_i(x) > 0$ for all x satisfying the constraints of (P). The present communication is concerned with the major generalization $(\mathcal{G}, \beta, \phi, h(\cdot, \cdot), \rho, \theta)$- univexity of the second order introduced by Verma (see (Verma, 2012)) that generalizes $(\mathcal{F}, \beta, \phi, \rho, \theta)$-univexity introduced by Zalmai (see (Zalmai, 2012)) and the first order univexity studied by Zalmai and Zhang (see (Zalmai & Zhang, 2007)) with its applications to parametric duality models in minimax fractional programming. The obtained results not only generalize the work of Zalmai on second order univexities, but also generalize other investigations on general invexities, including the valued-contributions of Jeyakumar (see (Jeyakumar *b*, 1985)), Liu (see (Liu, 1999)), Mangasarian (see (Mangasarian, 1975)), Mishra (see (Mishra, 1997), (Mishra, 2000)), Mishra and Rueda (see (Mishra & Rueda, 2000), (Mishra & Rueda, 2006)), Mond (see (Mond, 1974)) and others. Based on Mangasarian's second-order dual problem, Mond (see (Mond, 1974)) established some duality results under relatively simpler conditions involving a certain second-order generalization of the concept of convexity, while observed some possible computational advantages of second-order duality results, and also studied a pair of second-order symmetric dual problems. Mond's original notion of second-order convexity was followed by generalizations by other authors in different ways and applied establishing several second-order duality results for several classes of nonlinear programming problems. Although there exist various second-order duality results in the related literature for several classes of mathematical programming problems with a finite number of constraints, we feel our second-order duality results established in this paper are new and general in nature to the context of semiinfinite programming. For more details on second order duality results, we refer the reader (see (Aghezzaf, 2003) - (Zalmai & Zhang, 2007)), but more importantly, (see (Aghezzaf, 2003) - (Jeyakumar *b*, 1985), (Mond & Weir, 1981-1983), (Mond & Zhang, 1995) - (Zalmai & Zhang, 2007)).

Note that second-order duality for a conventional nonlinear programming problem is of the form

$$(P_0) \qquad \text{Minimize } f(x) \text{ subject to } g_i(x) \le 0, \quad i \in \underline{m}, \quad x \in \mathbb{R}^n,$$

where f and g_i, $i \in \underline{m}$, are twice differentiable real-valued functions defined on $\mathbb{R}^n$, was initially considered and studied by Mangasarian (see (Mangasarian, 1975)). The idea underlying his approach to constructing a second-order dual problem was based on taking linear and quadratic approximations of the objective and constraint functions about an arbitrary but fixed point, leading to the Wolfe dual of the approximated problem, and then allowing the fixed point to vary. Mangasarian (see (Mangasarian, 1975)), more specifically, formulated the following second-order dual problem for (P_0):

$$(D_0) \qquad \text{Maximize } f(y) + \sum_{i=1}^{m} u_i g_i(y) - \tfrac{1}{2}\left\langle z, \left[\nabla^2 f(y) + \sum_{i=1}^{m} u_i \nabla^2 g_i(y)\right] z\right\rangle$$

subject to

$$\nabla f(y) + \sum_{i=1}^{m} u_i \nabla g_i(y) + \Big[\nabla^2 f(y) + \sum_{i=1}^{m} u_i \nabla^2 g_i(y)\Big] z = 0,$$
$$y \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \quad u \geq 0, \quad z \in \mathbb{R}^n,$$

where $\nabla F(y)$ and $\nabla^2 F(y)$ denote, respectively, the gradient and Hessian of the function $F : \mathbb{R}^n \to \mathbb{R}$ evaluated at y and $\langle a, b\rangle$ denotes the inner product of the vectors a and b. Then, by imposing somewhat complicated conditions on f, g_i, $i \in \underline{m}$, and z, he proved weak, strong, and converse duality theorems for (P_0) and (D_0).

We observe that all the duality results established in this paper can easily be modified and restated for each one of the following classes of nonlinear programming problems, that are special cases of (P):

$(P1)$ Minimize $\dfrac{f_1(x)}{g_1(x)}$;

$(P2)$ Minimize $\max\limits_{1\leq i\leq p} f_i(x)$;

$(P3)$ $\underset{x\in\mathbb{F}}{\text{Minimize}}\ f_1(x)$,

where $\mathbb{F}$ (assumed to be nonempty) is the feasible set of (P), that is,

$$\mathbb{F} = \{x \in \mathbb{R}^n : G_j(x,t) \leq 0 \text{ for all } t \in T_j,\ j \in \underline{q}, \quad H_k(x,s) = 0 \text{ for all } s \in S_k,\ k \in \underline{r}\};$$

$(P4)$ Minimize $\max\limits_{1\leq i\leq p} \dfrac{f_i(x)}{g_i(x)}$

subject to

$$\tilde{G}_j(x) \leq 0, \quad j \in \underline{q}, \quad \tilde{H}_k(x) = 0, \quad k \in \underline{r}, \quad x \in \mathbb{R}^n,$$

where f_i and g_i, $i \in \underline{p}$, are as defined in the description of (P), and $\tilde{G}_j$, $j \in \underline{q}$, and $\tilde{H}_k$, $k \in \underline{r}$, are real-valued functions defined on X;

$(P5)$ $\underset{x\in\mathbb{G}}{\text{Minimize}}\ \dfrac{f_1(x)}{g_1(x)}$;

$(P6)$ $\underset{x\in\mathbb{G}}{\text{Minimize}}\ \max\limits_{1\leq i\leq p} f_i(x)$;

$(P7)$ $\underset{x\in\mathbb{G}}{\text{Minimize}}\ f_1(x)$,

where $\mathbb{G}$ is the feasible set of $(P4)$, that is,

$$\mathbb{G} = \{x \in \mathbb{R}^n : \tilde{G}_j(x) \leq 0,\ j \in \underline{q}, \quad \tilde{H}_k(x) = 0, \quad k \in \underline{r}\}.$$

2. Preliminaries

In this section we recall, the recently introduced major generalization $(\mathcal{G},\beta,\phi,h(\cdot,\cdot),\rho,\theta)-$ univexiity by Verma (see (Verma, 2012)) to the notion of the Zalmai type $(\mathcal{F},\beta,\phi,\rho,\theta)$-univexity of higher order (See (Zalmai, 2012)) to the context of parametric duality models in semiinfinite discrete minimax fractional programming. The obtained notion, in fact, reduces to most of the existing notions of invexities and univexities in the literature.

Recall that a function $\mathcal{G}:\mathbb{R}^n\to\mathbb{R}$ is said to be *sublinear*(*superlinear*) if

$$\mathcal{G}(x+y)\leq(\geq)\mathcal{G}(x)+\mathcal{G}(y)\;\forall\;x,y\in\mathbb{R}^n,$$

and $\mathcal{G}(ax)=a\mathcal{G}(x)$ for all $x\in\mathbb{R}^n$ and $a\in\mathbb{R}_+=[0,\infty)$.

Let $x^*\in X$ and let us assume that the function $f:X\to\mathbb{R}$ is twice continuously differentiable at x^*.

Definition 2.1. The function f is said to be *(strictly)* $(\mathcal{G},\beta,\phi,h(x^*,z),\rho,\theta)$-*univex* at x^* of higher order if there exist functions $\beta:X\times X\to\mathbb{R}_+\backslash\{0\}=(0,\infty)$, $\phi:\mathbb{R}\to\mathbb{R}$, $\rho:X\times X\to\mathbb{R}$, $\theta:X\times X\to\mathbb{R}^n$, and a sublinear function $\mathcal{G}(x,x^*;\cdot):\mathbb{R}^n\to\mathbb{R}$ such that for each $x\in X(x\neq x^*)$ and $z\in\mathbb{R}^n$,

$$\phi(f(x)-f(x^*)+\langle z,\nabla_z h(x^*,z)\rangle-h(x^*,z))(>)\geq\mathcal{G}(x,x^*;\beta(x,x^*)[\nabla_z h(x^*,z)]) + \rho(x,x^*)\|\theta(x,x^*)\|^2,$$

where $h:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ is differentiable with respect to the second component.

Definition 2.2. The function f is said to be *(strictly)* $(\mathcal{G},\beta,\phi,h(x^*,z),\rho,\theta)$-*pseudounivex* at x^* if there exist functions $\beta:X\times X\to\mathbb{R}_+\backslash\{0\}$, $\phi:\mathbb{R}\to\mathbb{R}$, $\rho:X\times X\to\mathbb{R}$, $\theta:X\times X\to\mathbb{R}^n$, and a sublinear function $\mathcal{G}(x,x^*;\cdot):\mathbb{R}^n\to\mathbb{R}$ such that for each $x\in X(x\neq x^*)$ and $z\in\mathbb{R}^n$,

$$\mathcal{G}(x,x^*;\beta(x,x^*)[\nabla_z h(x^*,z)])\geq-\rho(x,x^*)\|\theta(x,x^*)\|^2 \Rightarrow\quad\phi(f(x)-f(x^*)+\langle z,\nabla_z h(x^*,z)\rangle-h(x^*,z))(>)\geq 0,$$

equivalently,

$$\phi(f(x)-f(x^*)+\langle z,\nabla_z h(x^*,z)\rangle-h(x^*,z))(\leq)<0\Rightarrow \mathcal{G}(x,x^*;\beta(x,x^*)[\nabla_z h(x^*,z)])<-\rho(x,x^*)\|\theta(x,x^*)\|^2,$$

where $h:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ is differentiable with respect to the second component.

Definition 2.3. The function f is said to be *prestrictly* $(\mathcal{G},\beta,\phi,h(x^*,z),\rho,\theta)$-*pseudounivex* at x^* if there exist functions $\beta:X\times X\to\mathbb{R}_+\backslash\{0\}$, $\phi:\mathbb{R}\to\mathbb{R}$, $\rho:X\times X\to\mathbb{R}$, $\theta:X\times X\to\mathbb{R}^n$, and a sublinear function $\mathcal{G}(x,x^*;\cdot):\mathbb{R}^n\to\mathbb{R}$ such that for each $x\in X(x\neq x^*)$ and $z\in\mathbb{R}^n$,

$$\mathcal{G}(x,x^*;\beta(x,x^*)[\nabla_z h(x^*,z)])>-\rho(x,x^*)\|\theta(x,x^*)\|^2 \Rightarrow\quad\phi(f(x)-f(x^*)+\langle z,\nabla_z h(x^*,z)\rangle-h(x^*,z))\geq 0,$$

equivalently,

$$\phi(f(x) - f(x^*) + \langle z, \nabla_z h(x^*, z)\rangle - h(x^*, z)) < 0 \Rightarrow \\ \mathcal{G}(x, x^*; \beta(x, x^*)[\nabla_z h(x^*, z)]) \leq -\rho(x, x^*)\|\theta(x, x^*)\|^2,$$

where $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is differentiable with respect to the second component.

Definition 2.4. The function f is said to be *(prestrictly)*$(\mathcal{G}, \beta, \phi, h(x^*, z), \rho, \theta)$*-quasiunivex* at x^* if there exist functions $\beta : X \times X \to \mathbb{R}_+\backslash\{0\}$, $\phi : \mathbb{R} \to \mathbb{R}$, $\rho : X \times X \to \mathbb{R}$, $\theta : X \times X \to \mathbb{R}^n$, and a sublinear function $\mathcal{G}(x, x^*; \cdot) : \mathbb{R}^n \to \mathbb{R}$ such that for each $x \in X$ and $z \in \mathbb{R}^n$,

$$\phi(f(x) - f(x^*) + \langle z, \nabla_z h(x^*, z)\rangle - h(x^*, z))(<) \leq 0 \\ \Rightarrow \quad \mathcal{G}(x, x^*; \beta(x, x^*)[\nabla_z h(x^*, z)]) \leq -\rho(x, x^*)\|\theta(x, x^*)\|^2,$$

equivalently,

$$\mathcal{G}(x, x^*; \beta(x, x^*)[\nabla_z h(x^*, z)]) > -\rho(x, x^*)\|\theta(x, x^*)\|^2 \Rightarrow \\ \phi(f(x) - f(x^*) + \langle z, \nabla_z h(x^*, z)\rangle - h(x^*, z))(\geq) > 0,$$

where $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is differentiable with respect to the second component.

Definition 2.5. The function f is said to be *strictly* $(\mathcal{G}, \beta, \phi, h(x^*, z), \rho, \theta)$*-quasiunivex* at x^* if there exist functions $\beta : X \times X \to \mathbb{R}_+\backslash\{0\}$, $\phi : \mathbb{R} \to \mathbb{R}$, $\rho : X \times X \to \mathbb{R}$, $\theta : X \times X \to \mathbb{R}^n$, and a sublinear function $\mathcal{G}(x, x^*; \cdot) : \mathbb{R}^n \to \mathbb{R}$ such that for each $x \in X$ and $z \in \mathbb{R}^n$,

$$\phi(f(x) - f(x^*) + \langle z, \nabla_z h(x^*, z)\rangle - h(x^*, z)) \leq 0 \\ \Rightarrow \quad \mathcal{G}(x, x^*; \beta(x, x^*)[\nabla_z h(x^*, z)]) < -\rho(x, x^*)\|\theta(x, x^*)\|^2,$$

equivalently,

$$\mathcal{G}(x, x^*; \beta(x, x^*)[\nabla_z h(x^*, z)]) \geq -\rho(x, x^*)\|\theta(x, x^*)\|^2 \Rightarrow \\ \phi(f(x) - f(x^*) + \langle z, \nabla_z h(x^*, z)\rangle - h(x^*, z)) > 0,$$

where $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is differentiable with respect to the second component.

We note that the generalized $(\mathcal{G}, \beta, \phi, h(\cdot, \cdot), \rho, \theta)$-univexities (see (Verma, 2012)) at x^* of higher order reduce to the Zalmai type $(\mathcal{F}, \beta, \phi, \rho, \theta)$-univexities (see (Zalmai, 2012)) of higher-order if we set

$$h(x^*, z) = \langle z, \nabla f(x^*)\rangle + \frac{1}{2}\langle z, \nabla^2 f(x^*)z\rangle.$$

Then, we have

$$\nabla_z h(x^*, z) = \nabla f(x^*) + \nabla^2 f(x^*)z$$

and

$$\langle z, \nabla_z h(x^*, z)\rangle - h(x^*, z) = \frac{1}{2}\langle z, \nabla^2 f(x^*)z\rangle.$$

We observe some of the implications from the above definitions as follows: if f is $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-univex at x^*, then it is both $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-pseudounivex and $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-quasiunivex at x^*, if f is $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-quasiunivex at x^*, then it is prestrictly $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-quasiunivex at x^*, and if f is strictly $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-pseudounivex at x^*, then it is $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-quasiunivex at x^*.

Note that during the proofs of the duality theorems, sometimes it may be more convenient to use certain alternative but equivalent forms of the above definitions. We conclude this section by recalling a set of parametric necessary optimality conditions for (P) based on the following result.

Theorem 2.1. *(See (Verma, 2013)) Let $x^* \in \mathbb{F}$ and $\lambda^* = \max_{1\le i\le p} f_i(x^*)/g_i(x^*)$, for each $i \in \underline{p}$, let f_i and g_i be twice continuously differentiable at x^*, for each $j \in \underline{q}$, let the function $z \to G_j(z,t)$ be twice continuously differentiable at x^* for all $t \in T_j$, and for each $k \in \underline{r}$, let the function $z \to H_k(z,s)$ be twice continuously differentiable at x^* for all $s \in S_k$. If x^* is an optimal solution of (P), if the second order generalized Abadie constraint qualification holds at x^*, and if for any critical direction y, the set cone*

$$\begin{aligned} &\{\Big(\nabla G_j(x^*,t), \langle y, \nabla^2 G_j(x^*,t)y\rangle\Big) : t \in \hat{T}_j(x^*), j \in \underline{q}\} \\ + \quad &span\{\Big(\nabla H_k(x^*,s), \langle y, \nabla^2 H_k(x^*,s)y\rangle\Big) : s \in S_k, k \in \underline{r}\}, \end{aligned}$$

where $\hat{T}_j(x^*) = \{t \in T_j : G_j(x^*,t) = 0\}$, is closed, then there exist $u^* \in U = \{u \in \mathbb{R}^p : u \ge 0, \sum_{i=1}^p u_i = 1\}$ and integers ν_0^* and ν^*, with $0 \le \nu_0^* \le \nu^* \le n+1$, such that there exist ν_0^* indices j_m, with $1 \le j_m \le q$, together with ν_0^* points $t^m \in \hat{T}_{j_m}(x^*)$, $m \in \underline{\nu_0^*}$, $\nu^* - \nu_0^*$ indices k_m, with $1 \le k_m \le r$, together with $\nu^* - \nu_0^*$ points $s^m \in S_{k_m}$ for $m \in \underline{\nu^*}\backslash\underline{\nu_0^*}$, and ν^* real numbers v_m^*, with $v_m^* > 0$ for $m \in \underline{\nu_0^*}$, with the property that

$$\sum_{i=1}^{p} u_i^*[\nabla f_i(x^*) - \lambda^*(\nabla g_i(x^*)] + \sum_{m=1}^{\nu_0^*} v_m^*[\nabla G_{j_m}(x^*, t^m) + \sum_{m=\nu_0^*+1}^{\nu^*} v_m^* \nabla H_k(x^*, s^m) = 0, \tag{2.1}$$

$$\langle y, \Big[\sum_{i=1}^{p} u_i^*[\nabla^2 f_i(x^*) - \lambda^* \nabla^2 g_i(x^*)] + \sum_{m=1}^{\nu_0^*} v_m^* \nabla^2 G_{j_m}(x^*, t^m) + \sum_{m=\nu_0^*+1}^{\nu^*} v_m^* \nabla^2 H_k(x^*, s^m)\Big] y\rangle \ge 0. \tag{2.2}$$

We shall call x a *normal* feasible solution of (P) if x satisfies all the constraints of (P), if the generalized Abadie constraint qualification holds at x, and if the set $cone\{\nabla G_j(x,t) : t \in \hat{T}_j(x), j \in \underline{q}\} + span\{\nabla H_k(x,s) : s \in S_k, k \in \underline{r}\}$ is closed.

The above theorem on the necessary optimality conditions provides us with clear guidelines for formulating numerous Wolfe-type duality models for (P). From now on, the functions f_i, g_i, $i \in \underline{p}$, $z \to G_j(z,t)$, and $z \to H_k(z,s)$ are twice continuously differentiable on X for all $t \in T_j$, $j \in \underline{q}$, and all $s \in S_k$, $k \in \underline{r}$.

3. Duality Models

In this section, we consider two duality models with special constraint structures that allow for a greater variety of generalized $(\mathcal{G}, \beta, \phi, h(x,z), \rho, \theta)$-univexity conditions under which duality can be established based on the following set:

$$\mathbb{H} = \Big\{(y, z, u, v, \lambda, \nu, \nu_0, J_{\nu_0}, K_{\nu\setminus\nu_0}, \bar{t}, \bar{s}) : \ y \in X;\ z \in \mathbb{R}^n;\ u \in U;\ 0 \le \nu_0 \le \nu \le n+1;$$
$$v \in \mathbb{R}^{\nu}, v_i > 0, 1 \le i \le \nu_0;\ \lambda \in \mathbb{R}_+;\ J_{\nu_0} = (j_1, j_2, \ldots, j_{\nu_0}),\ 1 \le j_i \le q;\ K_{\nu\setminus\nu_0} =$$
$$(k_{\nu_0+1}, \ldots, k_{\nu}), 1 \le k_i \le r;\ \bar{t} = (t^1, t^2, \ldots, t^{\nu_0}),\ t^i \in T_{j_i};\ \bar{s} = (s^{\nu_0+1}, \ldots, s^{\nu}),\ s^i \in S_{k_i}\Big\}.$$

Consider the following two problems:

(DI) $\sup_{(y,z,u,v,\lambda,\nu,\nu_0,J_{\nu_0},K_{\nu\setminus\nu_0},\bar{t},\bar{s})\in\mathbb{H}} \lambda$

subject to

$$\sum_{i=1}^{p} u_i[\nabla_z h_i(y,z) - \lambda\nabla_z \kappa_i(y,z)] + \sum_{m=1}^{\nu_0} v_m[\nabla_z \mu_{j_m}(y, t^m, z) + \sum_{m=\nu_0+1}^{\nu} v_m[\nabla_z \psi_{k_m}(y, s^m, z)] = 0, \tag{3.1}$$

$$f_i(y) - \lambda g_i(y) + \sum_{i=1}^{p} u_i[h_i(y,z) - \lambda\kappa_i(y,z) - \langle z, \nabla_z h_i(y,z) - \lambda\nabla_z \kappa_i(y,z)]\rangle \ge 0, \quad i \in \underline{p}, \tag{3.2}$$

$$G_{j_m}(y, t^m) + \mu_{j_m}(y, t^m, z) - \langle z, \nabla_z \mu_{j_m}(y, t^m, z)\rangle \ge 0, \quad m \in \underline{\nu_0}, \tag{3.3}$$

$$v_m H_{k_m}(y, s^m) + v_m \psi_{k_m}(y, s^m, z) - \langle z, v_m \nabla_z \psi_{k_m}(y, s^m, z)\rangle \ge 0, \quad m \in \underline{\nu}\setminus\underline{\nu_0}; \tag{3.4}$$

($\tilde{DI}$) $\sup_{(y,z,u,v,\lambda,\nu,\nu_0,J_{\nu_0},K_{\nu\setminus\nu_0},\bar{t},\bar{s})\in\mathbb{H}} \lambda$

subject to (3.2)-(3.4) and

$$\mathcal{G}\Big(x, y; \sum_{i=1}^{p} u_i[\nabla_z h_i(y,z)] - \sum_{i=1}^{p} u_i\lambda[\nabla_z \kappa_i(y,z)] + \sum_{m=1}^{\nu_0} v_m[\nabla_z \mu_{j_m}(y, z, t^m)] + \sum_{m=\nu_0+1}^{\nu} v_m[\nabla_z \psi_{k_m}(y, z, s^m)] \ge 0 \ \text{ for all } x \in \mathbb{F}, \tag{3.5}$$

where $\mathcal{G}(x, y; \cdot)$ is a sublinear function from $\mathbb{R}^n$ to $\mathbb{R}$.

Note that if we Compare (DI) and ($\tilde{DI}$), we see that ($\tilde{DI}$) is relatively more general than (DI) in the sense that any feasible solution of (DI) is also feasible for ($\tilde{DI}$), but the converse may not be necessarily true.

Lemma 3.1. *(See (Zalmai, 2012)) For each $x \in X$,*

$$\varphi(x) \equiv \max_{1\le i\le p} \frac{f_i(x)}{g_i(x)} = \max_{u\in U} \frac{\sum_{i=1}^{p} u_i f_i(x)}{\sum_{i=1}^{p} u_i g_i(x)}.$$

The next theorem shows that (DI) is a dual problem for primal (P).

Theorem 3.1. *(Weak Duality) Let x and $w = (y, z, u, v, \lambda, \nu, \nu_0, J_{\nu_0}, K_{\nu\setminus\nu_0}, \bar{t}, \bar{s})$ be arbitrary feasible solutions of (P) and (DI), respectively, and let us assume that any one of the following five sets of hypotheses is satisfied:*

(a) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, \bar{\phi}, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \bar{\phi}, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \ge 0 \Rightarrow a \ge 0$;*

(ii) the function $\xi \to G_{j_m}(\xi, t^m)$ is $(\mathcal{G}, \beta, \hat{\phi}_m, \mu_m(\cdot,\cdot), \hat{\rho}_m, \theta)$-quasiunivex at y, $\hat{\phi}_m$ is increasing, and $\hat{\phi}_m(0) = 0$ for each $m \in \underline{\nu_0}$;

(iii) the function $\xi \to \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \breve{\phi}_m, \psi_m(\cdot,\cdot), \breve{\rho}_m, \theta)$- quasiunivex at y, $\breve{\phi}_m$ is increasing, and $\breve{\phi}_m(0) = 0$ for each $m \in \underline{\nu}\setminus\underline{\nu_0}$;

(iv) $\rho^*(x,y) + \sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m(x,y) + \sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x,y) \ge 0$ where $\rho^*(x,y) = \sum_{i=1}^{p} u_i[\bar{\rho}_i(x,y) + \lambda\tilde{\rho}_i(x,y)]$;

(b) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, \bar{\phi}, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \bar{\phi}, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \ge 0 \Rightarrow a \ge 0$;*

(ii) the function $\xi \to \sum_{m=1}^{\nu_0} \nu_m G_{j_m}(\xi, t^m)$ is $(\mathcal{G}, \beta, \hat{\phi}, \mu_m(\cdot,\cdot), \hat{\rho}, \theta)$-quasiunivex at y, $\hat{\phi}$ is increasing, and $\hat{\phi}(0) = 0$;

(iii) the function $\xi \to \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \breve{\phi}_m, \psi_m(\cdot,\cdot), \breve{\rho}_m, \theta)$-quasiunivex at y, $\breve{\phi}_m$ is increasing, and $\breve{\phi}_m(0) = 0$ for each $m \in \underline{\nu}\setminus\underline{\nu_0}$;

(iv) $\rho^*(x,y) + \hat{\rho}(x,y) + \sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x,y) \ge 0$;

(c) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, \bar{\phi}, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \bar{\phi}, \tilde{\rho}_i, \kappa_i(\cdot,\cdot), \theta)$-univex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \ge 0 \Rightarrow a \ge 0$;*

(ii) the function $\xi \to G_{j_m}(\xi, t^m)$ is $(\mathcal{G}, \beta, \hat{\phi}_m, \mu_m(\cdot,\cdot), \hat{\rho}_m, \theta)$-quasiunivex at y, $\hat{\phi}_m$ is increasing, and $\hat{\phi}_m(0) = 0$ for each $m \in \underline{\nu_0}$;

(iii) the function $\xi \to \sum_{m=\nu_0+1}^{\nu} \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \breve{\phi}, \psi_m(\cdot,\cdot), \breve{\rho}, \theta)$-quasiunivex at y, $\breve{\phi}$ is increasing, and $\breve{\phi}(0) = 0$;

(iv) $\rho^*(x,y) + \sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m(x,y) + \breve{\rho}(x,y) \ge 0$;

(d) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, \bar{\phi}, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \bar{\phi}, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \ge 0 \Rightarrow a \ge 0$;*

(ii) *the function* $\xi \to \sum_{m=1}^{\nu_0} v_m G_{j_m}(\xi, t^m)$ *is* $(\mathcal{G}, \beta, \hat{\phi}, \mu_m(\cdot,\cdot), \hat{\rho}, \theta)$*-quasiunivex at* y, $\hat{\phi}$ *is increasing, and* $\hat{\phi}(0) = 0$*;*

(iii) *the function* $\xi \to \sum_{m=\nu_0+1}^{\nu} v_m H_{k_m}(\xi, s^m)$ *is* $(\mathcal{G}, \beta, \breve{\phi}, \psi_m(\cdot,\cdot), \breve{\rho}, \theta)$*-quasiunivex at* y, $\breve{\phi}$ *is increasing, and* $\breve{\phi}(0) = 0$*;*

(iv) $\rho^*(x,y) + \hat{\rho}(x,y) + \breve{\rho}(x,y) \geq 0$*;*

(e) *(i)* *for each* $i \in \underline{p}$, f_i *is* $(\mathcal{G}, \beta, \bar{\phi}, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$*-univex and* $-g_i$ *is* $(\mathcal{G}, \beta, \bar{\phi}, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$*-univex at* y, $\bar{\phi}$ *is superlinear, and* $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$*;*

(ii) *the function* $\xi \to \sum_{m=1}^{\nu_0} v_m G_{j_m}(\xi, t^m) + \sum_{m=\nu_0+1}^{\nu} v_m H_{k_m}(\xi, s^m)$ *is* $(\mathcal{G}, \beta, \hat{\phi}, \tau_m, \hat{\rho}, \theta)$*-quasiunivex at* y, $\hat{\phi}$ *is increasing, and* $\hat{\phi}(0) = 0$*;*

(iii) $\rho^*(x,y) + \hat{\rho}(x,y) \geq 0$.

Then $\varphi(x) \geq \lambda$.

Proof. (a): Applying (i), we have the following inequality:

$$\begin{aligned}
&\bar{\phi}\Big(\sum_{i=1}^{p} u_i[f_i(x) - f_i(y)] + \Big\langle z, \sum_{i=1}^{p} u_i \nabla_z h_i(y,z) \Big\rangle - \sum_{i=1}^{p} u_i h_i(y,z) \\
&+ \lambda[\sum_{i=1}^{p} u_i[-g_i(x) + g_i(y)] - \Big\langle z, \sum_{i=1}^{p} u_i \nabla_z \kappa_i(y,z) \Big\rangle + \sum_{i=1}^{p} u_i \kappa_i(y,z)]\Big) \\
&\quad \geq \mathcal{G}\Big(x, y; \beta(x,y) \sum_{i=1}^{p} u_i\{\nabla_z h_i(y,z) - \lambda \nabla_z \kappa_i(y,z)\}\Big) + \sum_{i=1}^{p} u_i[\bar{\rho}_i(x,y) + \lambda \tilde{\rho}_i(x,y)]\|\theta(x,y)\|^2. \quad (3.6)
\end{aligned}$$

From the primal feasibility of x, dual feasibility of w, and (3.3), we find that

$$G_{j_m}(x, t^m) \leq 0 \leq G_{j_m}(y, t^m) + \mu_{j_m}(y, t^m, z) - \langle z, \nabla_z \mu_{j_m}(y, t^m, z)\rangle, \quad m \in \underline{\nu_0},$$

and hence using the properties of the functions $\hat{\phi}_m$, we have

$$\hat{\phi}_m(G_{j_m}(x, t^m) - [G_{j_m}(y, t^m) + \mu_{j_m}(y, t^m, z) - \langle z, \nabla_z \mu_{j_m}(y, t^m, z)\rangle]) \leq 0,$$

which from (ii) implies that $\mathcal{G}(x, y; \beta(x,y)[\nabla_z \mu_{j_m}(y, t^m, z)]) \leq -\hat{\rho}_m(x,y)\|\theta(x,y)\|^2$. As $v_m > 0$ for each $m \in \underline{\nu_0}$, the above inequality yield

$$\mathcal{G}\Big(x, y; \beta(x,y) \sum_{m=1}^{\nu_0} v_m[\langle z, \nabla_z \mu_{j_m}(y, t^m, z)\rangle]\Big) \leq -\sum_{m=1}^{\nu_0} v_m \hat{\rho}_m(x,y)\|\theta(x,y)\|^2. \quad (3.7)$$

Similarly, from the primal feasibility of x, dual feasibility of w, (3.4), and (iii) we deduce (since $v_m > 0$ for each $m \in \underline{\nu} \backslash \underline{\nu_0}$) that

$$\mathcal{G}\Big(x, y; \beta(x,y) \sum_{m=\nu_0+1}^{\nu} v_m[\nabla_z \psi_{j_m}(y, t^m, z)]\Big) \leq -\sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x,y)\|\theta(x,y)\|^2. \quad (3.8)$$

Now, based on the positivity of $\beta(x,y)$, sublinearity of $\mathcal{G}(x,y;\cdot)$, and (3.1), we conclude that

$$\mathcal{G}\Big(x,y;\beta(x,y)\sum_{i=1}^{p}u_i\{\nabla_z h_i(y,z)-\lambda\nabla_z\kappa_i(y,z)\}\Big)+\mathcal{G}(x,y;\beta(x,y)\sum_{m=1}^{\nu_0}v_m[\nabla_z\mu_{j_m}(y,t^m,z)])$$
$$+\quad \mathcal{G}\Big(x,y;\beta(x,y)\sum_{m=\nu_0+1}^{\nu}v_m[\nabla_z\psi_{j_m}(y,t^m,z)]\Big)\geq 0. \tag{3.9}$$

Next, applying (3.9) to (3.6), and then combining with (3.7) and (3.8) and using (iv), we have

$$\bar{\phi}\Big(\sum_{i=1}^{p}u_i[f_i(x)-f_i(y)]+\Big\langle z,\sum_{i=1}^{p}u_i\nabla_z h_i(y,z)\Big\rangle-\sum_{i=1}^{p}u_i h_i(y,z)$$
$$+\quad \lambda[\sum_{i=1}^{p}u_i[-g_i(x)+g_i(y)]-\Big\langle z,\sum_{i=1}^{p}u_i\nabla_z\kappa_i(y,z)\Big\rangle+\sum_{i=1}^{p}u_i\kappa_i(y,z)]\Big)$$
$$\geq\quad \mathcal{G}\Big(x,y;\beta(x,y)\sum_{i=1}^{p}u_i\{\nabla_z h_i(y,z)-\lambda\nabla_z\kappa_i(y,z)\}\Big)$$
$$+\quad \sum_{i=1}^{p}u_i[\bar{\rho}_i(x,y)+\lambda\tilde{\rho}_i(x,y)]\|\theta(x,y)\|^2\geq-\Big[\mathcal{G}(x,y;\beta(x,y)\sum_{m=1}^{\nu_0}v_m[\nabla_z\mu_{j_m}(y,t^m,z)])$$
$$+\quad \mathcal{G}\Big(x,y;\beta(x,y)\sum_{m=\nu_0+1}^{\nu}v_m[\nabla_z\psi_{j_m}(y,t^m,z)]\Big)\Big]\geq\sum_{m=1}^{\nu_0}v_m\hat{\rho}_m(x,y)\|\theta(x,y)\|^2+\sum_{m=\nu_0+1}^{\nu}\breve{\rho}_m(x,y)\|\theta(x,y)\|^2$$
$$+\quad \sum_{i=1}^{p}u_i[\bar{\rho}_i(x,y)+\lambda\tilde{\rho}_i(x,y)]\|\theta(x,y)\|^2=\sum_{m=1}^{\nu_0}v_m\hat{\rho}_m(x,y)\|\theta(x,y)\|^2$$
$$+\quad \sum_{m=\nu_0+1}^{\nu}\breve{\rho}_m(x,y)\|\theta(x,y)\|^2+\rho^*(x,y)\|\theta(x,y)\|^2\geq 0.$$

But $\bar{\phi}(a)\geq 0 \;\Rightarrow\; a\geq 0$ and hence because of (3.2) the above inequality reduces to

$$\sum_{i=1}^{p}u_i[f_i(x)-\lambda g_i(x)]\geq 0.$$

Finally, this inequality using Lemma 3.1 leads to the weak duality inequality as follows:

$$\varphi(x)=\max_{1\leq i\leq p}\frac{f_i(x)}{g_i(x)}=\max_{u\in U}\frac{\sum_{i=1}^{p}u_i f_i(x)}{\sum_{i=1}^{p}u_i g_i(x)}\geq\lambda.$$

(b) - (e) : The proofs are similar to that of part (a). □

The following theorem is based on the $(\mathcal{G},\beta,h_i(\cdot,\cdot),\tilde{\rho}_i,\theta)$-univexities and quasiunivexities.

Theorem 3.2. *(Weak Duality) Let x and $w=(y,z,u,v,\lambda,\nu,\nu_0,J_{\nu_0},K_{\nu\setminus\nu_0},\bar{t},\bar{s})$ be arbitrary feasible solutions of (P) and (DI), respectively, and let us assume that any one of the following five sets of hypotheses is satisfied:*

(a) (i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at y,

(ii) the function $\xi \to G_{j_m}(\xi, t^m)$ is $(\mathcal{G}, \beta, \mu_m(\cdot,\cdot), \hat{\rho}_m, \theta)$-quasiunivex at y, for each $m \in \underline{\nu_0}$;

(iii) the function $\xi \to \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \psi_m(\cdot,\cdot), \breve{\rho}_m, \theta)$-quasi univex at y, for each $m \in \underline{\nu} \backslash \underline{\nu_0}$;

(iv) $\rho^*(x,y) + \sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m(x,y) + \sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x,y) \geq 0$ where $\rho^*(x,y) = \sum_{i=1}^{p} u_i[\bar{\rho}_i(x,y) + \lambda \tilde{\rho}_i(x,y)]$;

(b) (i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$.

(ii) the function $\xi \to \sum_{m=1}^{\nu_0} \nu_m G_{j_m}(\xi, t^m)$ is $(\mathcal{G}, \beta, \mu_m(\cdot,\cdot), \hat{\rho}, \theta)$-quasiunivex at y.

(iii) the function $\xi \to \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \breve{\phi}_m, \psi_m(\cdot,\cdot), \breve{\rho}_m, \theta)$-quasiunivex at y.

(iv) $\rho^*(x,y) + \hat{\rho}(x,y) + \sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x,y) \geq 0$;

(c) (i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \tilde{\rho}_i, \kappa_i(\cdot,\cdot), \theta)$-univex at y.

(ii) the function $\xi \to G_{j_m}(\xi, t^m)$ is $(\mathcal{G}, \beta, \mu_m(\cdot,\cdot), \hat{\rho}_m, \theta)$-quasiunivex at y.

(iii) the function $\xi \to \sum_{m=\nu_0+1}^{\nu} \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \psi_m(\cdot,\cdot), \breve{\rho}, \theta)$-quasiunivex at y.

(iv) $\rho^*(x,y) + \sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m(x,y) + \breve{\rho}(x,y) \geq 0$;

(d) (i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at y.

(ii) the function $\xi \to \sum_{m=1}^{\nu_0} \nu_m G_{j_m}(\xi, t^m)$ is $(\mathcal{G}, \beta, \hat{\phi}, \mu_m(\cdot,\cdot), \hat{\rho}, \theta)$-quasiunivex at y.

(iii) the function $\xi \to \sum_{m=\nu_0+1}^{\nu} \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \breve{\phi}, \psi_m(\cdot,\cdot), \breve{\rho}, \theta)$-quasiunivex at y.

(iv) $\rho^*(x,y) + \hat{\rho}(x,y) + \breve{\rho}(x,y) \geq 0$;

(e) (i) for each $i \in \underline{p}$, f_i is $(\mathcal{G}, \beta, h_i(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex and $-g_i$ is $(\mathcal{G}, \beta, \kappa_i(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at y.

(ii) the function $\xi \to \sum_{m=1}^{\nu_0} \nu_m G_{j_m}(\xi, t^m) + \sum_{m=\nu_0+1}^{\nu} \nu_m H_{k_m}(\xi, s^m)$ is $(\mathcal{G}, \beta, \tau_m, \hat{\rho}, \theta)$-quasiunivex at y.

(iii) $\rho^*(x,y) + \hat{\rho}(x,y) \geq 0$.

Then $\varphi(x) \geq \lambda$.

Proof. (a): Applying (i), we have the following inequality:

$$\begin{aligned}
&\sum_{i=1}^{p} u_i[f_i(x) - f_i(y)] + \Big\langle z, \sum_{i=1}^{p} u_i \nabla_z h_i(y,z)\Big\rangle - \sum_{i=1}^{p} u_i h_i(y,z) \\
&+ \lambda[\sum_{i=1}^{p} u_i[-g_i(x) + g_i(y)] - \Big\langle z, \sum_{i=1}^{p} u_i \nabla_z \kappa_i(y,z)\Big\rangle + \sum_{i=1}^{p} u_i \kappa_i(y,z)] \\
&\geq \mathcal{G}\Big(x,y;\beta(x,y)\sum_{i=1}^{p} u_i\{\nabla_z h_i(y,z) - \lambda\nabla_z \kappa_i(y,z)\}\Big) + \sum_{i=1}^{p} u_i[\bar{\rho}_i(x,y) + \lambda\tilde{\rho}_i(x,y)]\|\theta(x,y)\|^2. \quad (3.10)
\end{aligned}$$

From the primal feasibility of x, dual feasibility of w, and (3.3), we find that

$$G_{j_m}(x,t^m) \leq 0 \leq G_{j_m}(y,t^m) + \mu_{j_m}(y,t^m,z) - \langle z, \nabla_z \mu_{j_m}(y,t^m,z)\rangle, \quad m \in \underline{\nu_0}.$$

Then we have $G_{j_m}(x,t^m) - [G_{j_m}(y,t^m) + \mu_{j_m}(y,t^m,z) - \langle z, \nabla_z \mu_{j_m}(y,t^m,z)\rangle] \leq 0$, which from (ii) implies that $\mathcal{G}(x,y;\beta(x,y)[\nabla_z \mu_{j_m}(y,t^m,z)]) \leq -\hat{\rho}_m(x,y)\|\theta(x,y)\|^2$. As $\nu_m > 0$ for each $m \in \underline{\nu_0}$, the above inequalities yield

$$\mathcal{G}\Big(x,y;\beta(x,y)\sum_{m=1}^{\nu_0} \nu_m[\langle z, \nabla_z \mu_{j_m}(y,t^m,z)\rangle]\Big) \leq -\sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m(x,y)\|\theta(x,y)\|^2. \quad (3.11)$$

Similarly, from the primal feasibility of x, dual feasibility of w, (3.4), and (iii) we deduce that

$$\mathcal{G}\Big(x,y;\beta(x,y)\sum_{m=\nu_0+1}^{\nu} \nu_m[\nabla_z \psi_{j_m}(y,t^m,z)]\Big) \leq -\sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x,y)\|\theta(x,y)\|^2. \quad (3.12)$$

Now, based on the positivity of $\beta(x,y)$, sublinearity of $\mathcal{G}(x,y;\cdot)$, and applying (3.1), we conclude that

$$\begin{aligned}
&\mathcal{G}\Big(x,y;\beta(x,y)\sum_{i=1}^{p} u_i\{\nabla_z h_i(y,z) - \lambda\nabla_z \kappa_i(y,z)\}\Big) + \mathcal{G}(x,y;\beta(x,y)\sum_{m=1}^{\nu_0} \nu_m[\nabla_z \mu_{j_m}(y,t^m,z)]) \\
+ \quad &\mathcal{G}\Big(x,y;\beta(x,y)\sum_{m=\nu_0+1}^{\nu} \nu_m[\nabla_z \psi_{j_m}(y,t^m,z)]\Big) \geq 0. \quad (3.13)
\end{aligned}$$

Next, applying (3.13) to (3.10), and then combining with (3.11) and (3.12) and using (iv), we have

$$\begin{aligned}
&\Big(\sum_{i=1}^{p} u_i[f_i(x) - f_i(y)] + \Big\langle z, \sum_{i=1}^{p} u_i \nabla_z h_i(y,z)\Big\rangle - \sum_{i=1}^{p} u_i h_i(y,z) \\
+ \quad &\lambda[\sum_{i=1}^{p} u_i[-g_i(x) + g_i(y)] - \Big\langle z, \sum_{i=1}^{p} u_i \nabla_z \kappa_i(y,z)\Big\rangle + \sum_{i=1}^{p} u_i \kappa_i(y,z)]\Big) \\
\geq \quad &\Big(\rho^*(x,y) + \sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m(x,y) + \sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x,y)\Big)\|\theta(x,y)\|^2 \geq 0.
\end{aligned}$$

Hence because of (3.2) the above inequality reduces to

$$\sum_{i=1}^{p} u_i[f_i(x) - \lambda g_i(x)] \geq 0.$$

Finally, this inequality using Lemma 3.1 leads to the weak duality inequality as follows:

$$\varphi(x) \equiv \max_{1\leq i\leq p} \frac{f_i(x)}{g_i(x)} = \max_{u\in U} \frac{\sum_{i=1}^{p} u_i f_i(x)}{\sum_{i=1}^{p} u_i g_i(x)} \geq \lambda.$$

(b) - (e) : The proofs are similar to that of part (a). □

Theorem 3.3. *(Strict Converse Duality) Let x^* be a normal optimal solution of (P), let $\tilde{w} = (\tilde{x}, \tilde{z}, \tilde{u}, \tilde{v}, \tilde{\lambda}, \tilde{\nu}, \tilde{\nu}_0, J_{\tilde{\nu}_0}, K_{\tilde{\nu}\setminus\tilde{\nu}_0}, \tilde{\bar{t}}, \tilde{\bar{s}})$ be an optimal solution of (DI), and assume that any one of the following five sets of conditions is satisfied:*

(a) The assumptions specified in part (a) of Theorem 3.2 are satisfied for the feasible solution $\tilde{w}$ of (DI). Moreover, $\bar{\phi}(a) > 0 \Rightarrow a > 0$, f_i is strictly $(\mathcal{G}, \beta, \bar{\phi}, h(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive, or $-g_i$ is strictly $(\mathcal{G}, \beta, \bar{\phi}, \kappa(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive (and $\tilde{\lambda} > 0$), or $\xi \to G_{j_m}(\xi, \tilde{t}^m)$ is strictly $(\mathcal{G}, \beta, \hat{\phi}_m, \mu(\cdot,\cdot), \hat{\rho}_m, \theta)$-pseudounivex at $\tilde{x}$ for at least one $m \in \underline{\tilde{\nu}_0}$, or $\xi \to \tilde{\nu}_m H_{k_m}(\xi, \tilde{s}^m)$ is strictly $(\mathcal{G}, \beta, \breve{\phi}_m, \psi(\cdot,\cdot), \breve{\rho}_m, \theta)$-pseudounivex at $\tilde{x}$ for at least one $m \in \underline{\tilde{\nu}}\setminus\underline{\tilde{\nu}_0}$, or

$$\rho^*(x^*, \tilde{x}) + \sum_{m=1}^{\tilde{\nu}_0} \tilde{\nu}_m \hat{\rho}_m(x^*, \tilde{x}) + \sum_{m=\tilde{\nu}_0+1}^{\tilde{\nu}} \tilde{\nu}_m \breve{\rho}_m(x^*, \tilde{x}) > 0,$$

where $\rho^*(x^*, \tilde{x}) = \sum_{i=1}^{p} \tilde{u}_i[\bar{\rho}_i(x^*, \tilde{x}) + \tilde{\lambda}\tilde{\rho}_i(x^*, \tilde{x})]$.

(b) The assumptions specified in part (b) of Theorem 3.2 are satisfied for the feasible solution $\tilde{w}$ of (DI). Moreover, $\bar{\phi}(a) > 0 \Rightarrow a > 0$, f_i is strictly $(\mathcal{G}, \beta, \bar{\phi}, h(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive, or $-g_i$ is strictly $(\mathcal{G}, \beta, \bar{\phi}, \kappa(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive (and $\tilde{\lambda} > 0$), or $\xi \to \sum_{m=1}^{\tilde{\nu}_0} \tilde{\nu}_m G_{j_m}(\xi, \tilde{t}^m)$ is strictly $(\mathcal{G}, \beta, \hat{\phi}, \mu(\cdot,\cdot), \hat{\rho}, \theta)$-pseudounivex at $\tilde{x}$, or $\xi \to \tilde{\nu}_m H_{k_m}(\xi, \tilde{s}^m)$ is strictly $(\mathcal{G}, \beta, \breve{\phi}_m, \psi(\cdot,\cdot), \breve{\rho}_m, \theta)$-pseudounivex at $\tilde{x}$ for at least one $m \in \underline{\tilde{\nu}}\setminus\underline{\tilde{\nu}_0}$, or $\rho^*(x^*, \tilde{x}) + \hat{\rho}(x^*, \tilde{x}) + \sum_{m=\tilde{\nu}_0+1}^{\tilde{\nu}} \tilde{\nu}_m \breve{\rho}_m(x^*, \tilde{x}) > 0$.

(c) The assumptions specified in part (c) of Theorem 3.2 are satisfied for the feasible solution $\tilde{w}$ of (DI). Moreover, $\bar{\phi}(a) > 0 \Rightarrow a > 0$, f_i is strictly $(\mathcal{G}, \beta, \bar{\phi}, h(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive, or $-g_i$ is strictly $(\mathcal{G}, \beta, \bar{\phi}, \kappa(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive (and $\tilde{\lambda} > 0$), or $\xi \to G_{j_m}(\xi, \tilde{t}^m)$ is strictly $(\mathcal{G}, \beta, \hat{\phi}_m, \mu(\cdot,\cdot), \hat{\rho}_m, \theta)$-pseudounivex at $\tilde{x}$ for at least one $m \in \underline{\tilde{\nu}_0}$, or $\xi \to \sum_{m=\tilde{\nu}_0+1}^{\tilde{\nu}} \tilde{\nu}_m H_{j_m}(\xi, \tilde{s}^m)$ is strictly $(\mathcal{G}, \beta, \breve{\phi}, \psi(\cdot,\cdot), \breve{\rho}, \theta)$-pseudounivex at $\tilde{x}$, or $\rho^*(x^*, \tilde{x}) + \sum_{m=1}^{\tilde{\nu}_0} \tilde{\nu}_m \hat{\rho}_m(x^*, \tilde{x}) + \tilde{\nu}_m \breve{\rho}(x^*, \tilde{x}) > 0$.

(d) The assumptions specified in part (d) of Theorem 3.2 are satisfied for the feasible solution $\tilde{w}$ of (DI). Moreover, $\bar{\phi}(a) > 0 \Rightarrow a > 0$, f_i is strictly $(\mathcal{G}, \beta, \bar{\phi}, h(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive, or $-g_i$ is strictly $(\mathcal{G}, \beta, \bar{\phi}, \kappa(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive (and $\tilde{\lambda} > 0$), or $\xi \to \sum_{m=1}^{\tilde{\nu}_0} \tilde{\nu}_m G_{j_m}(\xi, \tilde{t}^m)$ is strictly $(\mathcal{G}, \beta, \hat{\phi}, \mu(\cdot,\cdot), \hat{\rho}, \theta)$-pseudounivex at $\tilde{x}$, or $\xi \to \sum_{m=\tilde{\nu}_0+1}^{\tilde{\nu}} \tilde{\nu}_m H_{k_m}(\xi, \tilde{s}^m)$ is strictly $(\mathcal{G}, \beta, \breve{\phi}, \psi(\cdot,\cdot), \breve{\rho}_m, \theta)$-pseudounivex at $\tilde{x}$, or $\rho^*(x^*, \tilde{x}) + \hat{\rho}(x^*, \tilde{x}) + \breve{\rho}(x^*, \tilde{x}) > 0$.

(e) The assumptions specified in part (e) of Theorem 3.2 are satisfied for the feasible solution $\tilde{w}$ of (DI). Moreover, $\bar{\phi}(a) > 0 \Rightarrow a > 0$, f_i is strictly $(\mathcal{G}, \beta, \bar{\phi}, h(\cdot,\cdot), \bar{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive, or $-g_i$ is strictly $(\mathcal{G}, \beta, \bar{\phi}, \kappa(\cdot,\cdot), \tilde{\rho}_i, \theta)$-univex at $\tilde{x}$ for at least one $i \in \underline{p}$ with the corresponding component $\tilde{u}_i$ of $\tilde{u}$ positive (and $\tilde{\lambda} > 0$), or $\xi \to \sum_{m=1}^{\tilde{\nu}_0} \tilde{\nu}_m G_{j_m}(\xi, \tilde{t}^m) + \sum_{m=\tilde{\nu}_0+1}^{\tilde{\nu}} \tilde{\nu}_m H_{k_m}(\xi, \tilde{s}^m)$ is strictly $(\mathcal{G}, \beta, \hat{\phi}, \tau(\cdot,\cdot), \hat{\rho}, \theta)$-pseudounivex at $\tilde{x}$, or $\rho^*(x^*, \tilde{x}) + \hat{\rho}(x^*, \tilde{x}) > 0$.

Then $\tilde{x} = x^*$ and $\varphi(x^*) = \tilde{\lambda}$.

Proof. The proof is similar to that of Theorem 3.2. □

4. Specialization I

In this section, we consider two duality models with special constraint structures that allow the generalized $(\mathcal{G}, \beta, \phi, h(\cdot,\cdot), \rho, \theta)$-univexity reduce to second order generalized $(\mathcal{F}, \beta, \phi, \rho, \theta)$-univexity introduced and studied by Zalmai (see (Zalmai, 2012)) under which duality can be established.

Consider the following two problems:

(DII) $\sup_{(y,z,u,v,\lambda,\nu,\nu_0,J_{\nu_0},K_{\nu\setminus\nu_0},\bar{t},\bar{s})\in\mathbb{H}} \lambda$

subject to

$$\sum_{i=1}^{p} u_i[\nabla f_i(y) - \lambda \nabla g_i(y)] + \sum_{m=1}^{\nu_0} v_m \nabla G_{j_m}(y, t^m) + \sum_{m=\nu_0+1}^{\nu} v_m \nabla H_{k_m}(y, s^m)$$
$$+ \Big\{ \sum_{i=1}^{p} u_i[\nabla^2 f_i(y) - \lambda \nabla^2 g_i(y)] + \sum_{m=1}^{\nu_0} v_m \nabla^2 G_{j_m}(y, t^m) + \sum_{m=\nu_0+1}^{\nu} v_m \nabla^2 H_{k_m}(y, s^m) \Big\} z = 0, \quad (4.1)$$

$$f_i(y) - \lambda g_i(y) - \frac{1}{2}\langle z, [\nabla^2 f_i(y) - \lambda \nabla^2 g_i(y)] z\rangle \geq 0, \quad i \in \underline{p}, \quad (4.2)$$

$$G_{j_m}(y, t^m) - \frac{1}{2}\langle z, \nabla^2 G_{j_m}(y, t^m) z\rangle \geq 0, \quad m \in \underline{\nu_0}, \quad (4.3)$$

$$v_m H_{k_m}(y, s^m) - \frac{1}{2}\langle z, v_m \nabla^2 H_{k_m}(y, s^m) z\rangle \geq 0, \quad m \in \underline{\nu}\setminus\underline{\nu_0}; \quad (4.4)$$

$$(\tilde{D}II) \qquad \sup_{(y,z,u,v,\lambda,\nu,\nu_0,J_{\nu_0},K_{\nu\setminus\nu_0},\bar{t},\bar{s})\in\mathbb{H}} \lambda \text{ subject to (3.3) and (4.2) - (4.4).}$$

The next theorem shows that (DII) is a dual problem for (P).

Theorem 4.1. *(Weak Duality) Let x and $w = (y, z, u, v, \lambda, \nu, \nu_0, J_{\nu_0}, K_{\nu\setminus\nu_0}, \bar{t}, \bar{s})$ be arbitrary feasible solutions of (P) and (DII), respectively, and assume that any one of the following five sets of hypotheses is satisfied:*

(a) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$-sounivex and $-g_i$ is $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$-sounivex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$;*

(ii) the function $\xi \to G_{j_m}(\xi, t^m)$ is $(\mathcal{F}, \beta, \hat{\phi}_m, \hat{\rho}_m, \theta)$-quasisounivex at y, $\hat{\phi}_m$ is increasing, and $\hat{\phi}_m(0) = 0$ for each $m \in \underline{\nu_0}$;

(iii) the function $\xi \to v_m H_{k_m}(\xi, s^m)$ is $(\mathcal{F}, \beta, \breve{\phi}_m, \breve{\rho}_m, \theta)$-quasisounivex at y, $\breve{\phi}_m$ is increasing, and $\breve{\phi}_m(0) = 0$ for each $m \in \underline{\nu}\setminus\underline{\nu_0}$;

(iv) $\rho^*(x, y) + \sum_{m=1}^{\nu_0} v_m\hat{\rho}_m(x, y) + \sum_{m=\nu_0+1}^{\nu} v_m\breve{\rho}_m(x, y) \geq 0$, where $\rho^*(x, y) = \sum_{i=1}^{p} u_i[\bar{\rho}_i(x, y) + \lambda\tilde{\rho}_i(x, y)]$;

(b) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$-sounivex and $-g_i$ is $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$-sounivex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$;*

(ii) the function $\xi \to \sum_{m=1}^{\nu_0} v_m G_{j_m}(\xi, t^m)$ is $(\mathcal{F}, \beta, \hat{\phi}, \hat{\rho}, \theta)$-quasisounivex at y, $\hat{\phi}$ is increasing, and $\hat{\phi}(0) = 0$;

(iii) the function $\xi \to v_m H_{k_m}(\xi, s^m)$ is $(\mathcal{F}, \beta, \breve{\phi}_m, \breve{\rho}_m, \theta)$-quasisounivex at y, $\breve{\phi}_m$ is increasing, and $\breve{\phi}_m(0) = 0$ for each $m \in \underline{\nu}\setminus\underline{\nu_0}$;

(iv) $\rho^*(x, y) + \hat{\rho}(x, y) + \sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m(x, y) \geq 0$;

(c) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$-sounivex and $-g_i$ is $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$-sounivex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$;*

(ii) the function $\xi \to G_{j_m}(\xi, t^m)$ is $(\mathcal{F}, \beta, \hat{\phi}_m, \hat{\rho}_m, \theta)$-quasisounivex at y, $\hat{\phi}_m$ is increasing, and $\hat{\phi}_m(0) = 0$ for each $m \in \underline{\nu_0}$;

(iii) the function $\xi \to \sum_{m=\nu_0+1}^{\nu} v_m H_{k_m}(\xi, s^m)$ is $(\mathcal{F}, \beta, \breve{\phi}, \breve{\rho}, \theta)$-quasisounivex at y, $\breve{\phi}$ is increasing, and $\breve{\phi}(0) = 0$;

(iv) $\rho^*(x, y) + \sum_{m=1}^{\nu_0} v_m\hat{\rho}_m(x, y) + \breve{\rho}(x, y) \geq 0$;

(d) *(i) for each $i \in \underline{p}$, f_i is $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$-sounivex and $-g_i$ is $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$-sounivex at y, $\bar{\phi}$ is superlinear, and $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$;*

(ii) the function $\xi \to \sum_{m=1}^{\nu_0} v_m G_{j_m}(\xi, t^m)$ is $(\mathcal{F}, \beta, \hat{\phi}, \hat{\rho}, \theta)$-quasisounivex at y, $\hat{\phi}$ is increasing, and $\hat{\phi}(0) = 0$;

(iii) *the function* $\xi \to \sum_{m=\nu_0+1}^{\nu} v_m H_{k_m}(\xi, s^m)$ *is* $(\mathcal{F}, \beta, \breve{\phi}, \breve{\rho}, \theta)$*-quasisounivex at* y, $\breve{\phi}$ *is increasing, and* $\breve{\phi}(0) = 0$*;*

(iv) $\rho^*(x, y) + \hat{\rho}(x, y) + \breve{\rho}(x, y) \geq 0$*;*

(e) *(i) for each* $i \in \underline{p}$, f_i *is* $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$*-sounivex and* $-g_i$ *is* $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$*-sounivex at* y, $\bar{\phi}$ *is superlinear, and* $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$*;*

(ii) the function $\xi \to \sum_{m=1}^{\nu_0} v_m G_{j_m}(\xi, t^m) + \sum_{m=\nu_0+1}^{\nu} v_m H_{k_m}(\xi, s^m)$ *is* $(\mathcal{F}, \beta, \hat{\phi}, \hat{\rho}, \theta)$*-quasisounivex at* y, $\hat{\phi}$ *is increasing, and* $\hat{\phi}(0) = 0$*;*

(iii) $\rho^*(x, y) + \hat{\rho}(x, y) \geq 0$.

Then $\varphi(x) \geq \lambda$.

Proof. The poof is similar to that of Theorem 3.2. □

5. Specializations II

In this section, we consider certain specializations of the $(\mathcal{G}, \beta, \phi, h(\cdot, \cdot), \rho, \theta)$-univexity to first order univexity under which first order duality (see (Zalmai & Zhang, 2007)) can be established. These duality models have the following forms:

$(DIII)$ $\sup_{(y,u,v,\lambda,\nu,\nu_0,J_{\nu_0},K_{\nu\setminus\nu_0},\bar{t},\bar{s})\in\mathbb{H}} \lambda$

subject to

$$\sum_{i=1}^{p} u_i[\nabla f_i(y) - \lambda \nabla g_i(y)] + \sum_{m=1}^{\nu_0} v_m \nabla G_{j_m}(y, t^m) + \sum_{m=\nu_0+1}^{\nu} v_m \nabla H_{k_m}(y, s^m) = 0, \tag{5.1}$$

$$u_i[f_i(y) - \lambda g_i(y)] \geq 0, \quad i \in \underline{p}, \tag{5.2}$$

$$G_{j_m}(y, t^m) \geq 0, \quad m \in \underline{\nu_0}, \tag{5.3}$$

$$v_m H_{k_m}(y, s^m) \geq 0, \quad m \in \underline{\nu}\setminus\underline{\nu_0}; \tag{5.4}$$

$(\tilde{D}III)$ $\sup_{(y,u,v,\lambda,\nu,\nu_0,J_{\nu_0},K_{\nu\setminus\nu_0},\bar{t},\bar{s})\in\mathbb{H}} \lambda$

subject to (3.3) and (5.2) - (5.4).

Theorem 5.1. *(see (Zalmai & Zhang, 2007)) (Weak Duality) Let* x *and* $(y, u, v, \lambda, \nu, \nu_0, J_{\nu_0}, K_{\nu\setminus\nu_0}, \bar{t}, \bar{s})$ *be arbitrary feasible solutions of (P) and (DIII), respectively, and assume that any one of the following five sets of hypotheses is satisfied:*

(a) *(i) for each* $i \in \underline{p}$, f_i *is* $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$*-univex and* $-g_i$ *is* $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$*-univex at* y, $\bar{\phi}$ *is superlinear, and* $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$*;*

(ii) *the function* $z \to G_{j_m}(z, t^m)$ *is* $(\mathcal{F}, \beta, \hat{\phi}_m, \hat{\rho}_m, \theta)$*-quasiunivex at* y, $\hat{\phi}_m$ *is increasing, and* $\hat{\phi}_m(0) = 0$ *for each* $m \in \underline{\nu_0}$*;*

(iii) *the function* $z \to \nu_m H_{k_m}(z, s^m)$ *is* $(\mathcal{F}, \beta, \breve{\phi}_m, \breve{\rho}_m, \theta)$*-quasiunivex at* y, $\breve{\phi}_m$ *is increasing, and* $\breve{\phi}_m(0) = 0$ *for each* $m \in \underline{\nu} \backslash \underline{\nu_0}$*;*

(iv) $\rho^* + \sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m + \sum_{m=\nu_0+1}^{\nu} \nu_m \breve{\rho}_m \geq 0$*, where* $\rho^* = \sum_{i=1}^{p} u_i(\bar{\rho}_i + \lambda \tilde{\rho}_i)$*;*

(b) *(i)* *for each* $i \in \underline{p}$, f_i *is* $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$*-univex and* $-g_i$ *is* $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$*-univex at* y, $\bar{\phi}$ *is superlinear, and* $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$*;*

(ii) *the function* $z \to \sum_{m=1}^{\nu_0} \nu_m G_{j_m}(z, t^m)$ *is* $(\mathcal{F}, \beta, \hat{\phi}, \hat{\rho}, \theta)$*-quasiunivex at* y, $\hat{\phi}$ *is increasing, and* $\hat{\phi}(0) = 0$*;*

(iii) *the function* $z \to \nu_m H_{k_m}(z, s^m)$ *is* $(\mathcal{F}, \beta, \breve{\phi}_m, \breve{\rho}_m, \theta)$*-quasiunivex at* y, $\breve{\phi}_m$ *is increasing, and* $\breve{\phi}_m(0) = 0$ *for each* $m \in \underline{\nu} \backslash \underline{\nu_0}$*;*

(iv) $\rho^* + \hat{\rho} + \sum_{m=\nu_0+1}^{\nu} \breve{\rho}_m \geq 0$*;*

(c) *(i)* *for each* $i \in \underline{p}$, f_i *is* $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$*-univex and* $-g_i$ *is* $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$*-univex at* y, $\bar{\phi}$ *is superlinear, and* $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$*;*

(ii) *the function* $z \to G_{j_m}(z, t^m)$ *is* $(\mathcal{F}, \beta, \hat{\phi}_m, \hat{\rho}_m, \theta)$*-quasiunivex at* y, $\hat{\phi}_m$ *is increasing, and* $\hat{\phi}_m(0) = 0$ *for each* $m \in \underline{\nu_0}$*;*

(iii) *the function* $z \to \sum_{m=\nu_0+1}^{\nu} \nu_m H_{k_m}(z, s^m)$ *is* $(\mathcal{F}, \beta, \breve{\phi}, \breve{\rho}, \theta)$*-quasiunivex at* y, $\breve{\phi}$ *is increasing, and* $\breve{\phi}(0) = 0$*;*

(iv) $\rho^* + \sum_{m=1}^{\nu_0} \nu_m \hat{\rho}_m + \breve{\rho} \geq 0$*;*

(d) *(i)* *for each* $i \in \underline{p}$, f_i *is* $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$*-univex and* $-g_i$ *is* $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$*-univex at* y, $\bar{\phi}$ *is superlinear, and* $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$*;*

(ii) *the function* $z \to \sum_{m=1}^{\nu_0} \nu_m G_{j_m}(z, t^m)$ *is* $(\mathcal{F}, \beta, \hat{\phi}, \hat{\rho}, \theta)$*-quasiunivex at* y, $\hat{\phi}$ *is increasing, and* $\hat{\phi}(0) = 0$*;*

(iii) *the function* $z \to \sum_{m=\nu_0+1}^{\nu} \nu_m H_{k_m}(z, s^m)$ *is* $(\mathcal{F}, \beta, \breve{\phi}, \breve{\rho}, \theta)$*-quasiunivex at* y, $\breve{\phi}$ *is increasing, and* $\breve{\phi}(0) = 0$*;*

(iv) $\rho^* + \hat{\rho} + \breve{\rho} \geq 0$*;*

(e) *(i)* *for each* $i \in \underline{p}$, f_i *is* $(\mathcal{F}, \beta, \bar{\phi}, \bar{\rho}_i, \theta)$*-univex and* $-g_i$ *is* $(\mathcal{F}, \beta, \bar{\phi}, \tilde{\rho}_i, \theta)$*-univex at* y, $\bar{\phi}$ *is superlinear, and* $\bar{\phi}(a) \geq 0 \Rightarrow a \geq 0$*;*

(ii) *the function* $z \to \sum_{m=1}^{\nu_0} \nu_m G_{j_m}(z, t^m) + \sum_{m=\nu_0+1}^{\nu} \nu_m H_{k_m}(z, s^m)$ *is* $(\mathcal{F}, \beta, \hat{\phi}, \hat{\rho}, \theta)$*-quasiunivex at* y, $\hat{\phi}$ *is increasing, and* $\hat{\phi}(0) = 0$*;*

(iii) $\rho^* + \hat{\rho} \geq 0$*.*

Then $\varphi(x) \geq \lambda$*.*

6. Concluding Remarks

The duality results established in this communication encompass a fairly large number of second-order dual problems and duality theorems that were investigated previously for several classes of nonlinear programming problems. Furthermore, the methods utilized in this paper could lead to extend and generalize results to other classes of mathematical programming problems based on general univexity assumptions.

Acknowledgment

The author is greatly indebted to the reviewer for all highly constructive comments and valuable suggestions leading to the revised version.

References

Aghezzaf, B. (2003). Second order mixed type duality in multiobjective programming problems, *J. Math. Anal. Appl.* **285**, 97–106.

Ahmad, I. and Z. Husain (2005). Nondifferentiable second-order symmetric duality, *Asia-Pacific J. Oper. Res.* **22**, 19–31.

Ahmad, I., Z. Husain and S. Sharma (2007). Higher-order duality in nondifferentiable multiobjective programming, *Numer. Func. Anal. Optim.* **28**, 989–1002.

Ahmad, I. and S. Sharma (2007). Second-order duality for nondifferentiable multiobjective programming problems, *Numer. Func. Anal. Optim.* **28**, 975–988.

Bector, C. R. and B. K. Bector (1986). Generalized bonvex functions and second-order duality for a nonlinear programming problem, *Congressus Numer.* **22**, 37–52.

Bector, C. R. and B. K. Bector (1986). On various duality theorems for second-order duality in nonlinear programming, *Cahiers du Centre d'Etudes de Rechereche Opér.* **28**, 283–292.

Bector, C. R. and S. Chandra (1986). Second-order duality for generalized fractional programming, *Methods Oper. Res.* **56**, 11–28.

Bector, C. R. and S. Chandra (1986). Second order symmetric and self dual programs, *Opsearch* **23**, 89–95.

Bector, C. R. and S. Chandra (1986). First and second order duality for a class of nondifferentiable fractional programming problems, *J. Inform. Optim. Sci.* **7**, 335–348.

Bector, C. R. and S. Chandra (1987). Generalized bonvexity and higher order duality for fractional programming, *Opsearch* **24**, 143–154.

Bector, C. R., S. Chandra and I. Husain (1991). Second-order duality for a minimax programming problem, *Opsearch* **28**, 249–263.

Chen, X. (2008). Sufficient conditions and duality for a class of multiobjective fractional programming problems with higher-order (F, α, ρ, d)-convexity, *J. Appl. Math. Comput.* **28**, 107–121.

Egudo, R. R. and M. A. Hanson (1993). Second order duality in multiobjective programming, *Opsearch* **30**, 223–230.

Gulati, T. R. and D. Agarwal (2007). Second-order duality in multiobjective programming involving (F, α, ρ, d)-V-type I functions, *Numer. Funct. Anal. Optim.* **28**, 1263–1277.

Gulati, T. R. and D. Agarwal (2007). On Huard type second-order converse duality in nonlinear programming, *Appl. Math. Lett.* **20**, 1057–1063.

Gulati, T. R. and D. Agarwal (2008). Optimality and duality in nondifferentiable multiobjective mathematical programming involving higher order (F, α, ρ, d)-type I functions, *J. Appl. Comput.* **27**, 345–364.

Gulati, T. R. and I. Ahmad (1997). Second order symmetric duality for nonlinear minimax mixed integer programming problems, *European J. Oper. Res.* **101**, 122–129.

Gulati, T.R., I. Ahmad and I. Husain (2001). Second order symmetric duality with generalized convexity, *Opsearch* **38**, 210–222.

Gulati, T. R. and Geeta (2010). Mond-Weir type second-order symmetric duality in multiobjective programming over cones, *Appl. Math. Lett.* **23**, 466–471.

Gulati, T. R. and S. K. Gupta (2007). Second-order symmetric duality for minimax integer programs over cones, *Internat. J. Oper. Res.* **4**, 181–188.

Gulati, T. R. and S. K. Gupta (2007). Higher-order nondifferentiable symmetric duality with generalized F-convexity, *J. Math. Anal. Appl.* **329**, 229–237.

Gulati, T. R. and S. K. Gupta (2007). A note on Mond-Weir type second-order symmetric duality, *Asia-Pac. J. Oper. Res.* **24**, 737–740.

Gulati, T. R. and S. K. Gupta (2009). Higher-order symmetric duality with cone constraints, *Appl. Math. Lett.* **22**, 776–781.

Gulati, T. R., S. K. Gupta and I. Ahmad (2008). Second-order symmetric duality with cone constraints, *J. Comput. Appl. Math.* **220**, 347–354.

Gulati, T. R. and G. Mehndiratta (2010). Nondifferentiable multiobjective Mond-Weir type second-order symmetric duality over cones, *Optim. Lett.* 4, 293–309.

Gulati, T. R., H. Saini and S. K. Gupta (2010). Second-order multiobjective symmetric duality with cone constraints, *European J. Oper. Res.* **205**, 247–252.

Gupta, S. K. and N. Kailey (2010). A note on multiobjective second-order symmetric duality, *European J. Oper. Res.* **201**, 649–651.

Hachimi, M. and B. Aghezzaf (2004). Second order duality in multiobjective programming involving generalized type-I functions, *Numer. Funct. Anal. Optim.* **25**, 725–736.

Hanson, M. A. (1993). Second order invexity and duality in mathematical programming, *Opsearch* **30**, 313–320.

Hou, S. H. and X. M. Yang (2002). On second-order symmetric duality in nondifferentiable programming, *J. Math. Anal. Appl.* **255**, 491–498.

Husain, Z., I. Ahmad and S. Sharma (2009). Second order duality for minmax fractional programming, *Optim. Lett.* **3**, 277–286.

Husain, I., A. Goyel and M. Masoodi (2007). Second order symmetric and maxmin symmetric duality with cone constraints, *Internat. J. Oper Res.* **4**, 199–205.

Husain, I. and Z. Jabeen (2004). Second order duality for fractional programming with support functions, *Opsearch* **41**, 121–134.

Jeyakumar, V. (1985). ρ-Convexity and second order duality, *Utilitas Math.* **29**, 71–85.

Jeyakumar, V. (1985). First and second order fractional programming duality, *Opsearch* **22**, 24–41.

Liu, J. C. (1999). Second order duality for minimax programming, *Utilitas Math.* **56**, 53–63

Mangasarian, O. L. (1975). Second- and higher-order duality theorems in nonlinear programming, *J. Math. Anal. Appl.* **51**, 607–620.

Mishra, S. K. (1997). Second order generalized invexity and duality in mathematical programming, *Optimization* **42**, 51–69.

Mishra, S. K. (2000). Second order symmetric duality in mathematical programming with F-convexity, *European J. Oper. Res.* **127**, 507–518.

Mishra, S. K. and N. G. Rueda (2000). Higher-order generalized invexity and duality in mathematical programming, *J. Math. Anal. Appl.* **247**, 173–182.

Mishra, S. K. and N. G. Rueda (2006). Second-order duality for nondifferentiale minimax programming involving generalized type I functions, *J. Optim. Theory Appl.* **130**, 477–486.

Mond, B. (1974). Second order duality for nonlinear programs, *Opsearch* **11**, 90–99.

Mond, B. and T. Weir (1981–1983). Generalized convexity and higher-order duality, *J. Math. Sci.* **16-18**, 74–94.

Mond, B. and T. Weir (1981). Generalized concavity and duality, in *Generalized Concavity in Optimization and Economics* (S. Schaible and W. T. Ziemba, eds.), *Academic Press, New York,* 1981, pp. 263–279.

Mond, B. and J. Zhang (1995). Duality for multiobjective programming involving second-order V-invex functions, in *Proceedings of the Optimization Miniconference II* (B. M. Glover and V. Jeyakumar, eds.), University of New South Wales, Sydney, Australia, 1995, pp. 89–100.

Mond, B. and J. Zhang (1998). *Higher order invexity and duality in mathemaical programming*, in Generalized Convexity, Generalized Monotonicity : Recent Results (J. P. Crouzeix, *et al.*, eds.), Kluwer Academic Publishers, printed in the Netherlands, 1998, pp. 357–372.

Patel, R. B. (1997). Second order duality in multiobjective fractional programming, *Indian J. Math.* **38**, 39–46.

Srivastava, M. K. and M. Bhatia (2006). Symmetric duality for multiobjective programming using second order (F,ρ)-convexity, *Opsearch* **43**, 274–295.

Srivastava, S. K. and M. G. Govil (2000). Second order duality for multiobjective programming involving (F,ρ,σ)-type I functions, *Opsearch* **37**, 316–326.

Suneja, C. S. K., S. Lalitha and S. Khurana (2003). Second order symmetric duality in multiobjective programming, *European J. Oper. Res.* **144**, 492–500.

Suneja, S. K., M. K. Srivastava and M. Bhatia (2008). Higher order duality in multiobjective fractional programming with support functions, *J. Math. Anal. Appl.* **347**, 8–17.

Verma, R. U. (2012). A generalization to Zalmai type univexities and applications to parametric Duality models in discrete minimax fractional programming, *Advances in Nonlinear Variational Inequalities* **15**(2), 113–123.

Verma, R. U. (2013). Generalized $(\mathcal{G},b,\beta,\phi,h,\rho,\theta)$-univexities with applications to parametric duality models for discrete minimax fractional programming, *Transactions on Mathematical Programming and Applications* **1**(1), 1–14.

Zalmai, G. J. (2012).Generalized second-order $(\mathcal{F},\beta,\phi,\rho,\theta)-$univex functions and parametric duality models in semi-infinite discrete minmax fractional programming, *Advances in Nonlinear Variational Inequalities* **15**(2), 63–91.

Zalmai, G. J. (1999). Optimality conditions and duality for constrained measurable subset selection problems with minmax objective functions, *Optimization* **2**, 377–395.

Zalmai, G. J. and Q. Zhang (2007). Generalized $(\mathcal{F},\beta,\phi,\rho,\theta)$-univex functions and parametric duality models in semiinfinite discrete minmax fractional programming, *Advances in Nonlinear Variational Inequalities* **10**, 21–42.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 85–89

On Univalent Functions with Logarithmic Coefficients by Using Convolution

Sh. Najafzadeh[a,*], A. Ebadian[a]

[a]*Department of Mathematics, Payame Noor University, Iran*

Abstract

The purpose of this present paper is to derive some inclusion results and coefficient estimates for certain analytic functions with logarithmic coefficients by using Hadamard product. Relevant connections of the results with various known properties are also investigated.

Keywords: Convolution, univalent function, logarithmic coefficient, coefficient bounds, extreme point, convex set.
2010 MSC: 30C45, 30C50.

1. Introduction and Motivation

let A denote the class of normalized functions $f(z)$ of the from

$$f(z) = z + \sum_{n=2}^{+\infty} a_n z^n, \tag{1.1}$$

which are holomorphic in the open unit disk $\Delta = \{z : |z| < 1\}$. Let N denote the subclass of A consisting of functions $f(z)$ of the form

$$f(z) = z - \sum_{n=2}^{+\infty} a_n z^n. \quad (a_n \geq 0). \tag{1.2}$$

Associated with each f in A is a well defined logarithmic function

$$\log \frac{f(z)}{z} = 2 \sum_{n=1}^{+\infty} \gamma_n z^n. \ z \in \Delta. \tag{1.3}$$

*Corresponding author
Email addresses: najafzadeh1234@yahoo.ie (Sh. Najafzadeh), aebadian@yahoo.com (A. Ebadian)

The numbers γ_n are called the logarithmic coefficients of $f(z)$. See (Girela, 2000). For $\log \frac{f(z)}{z}$ given by (1.3) and $G(z) \in N$ given by

$$G(z) = z - \sum_{n=2}^{+\infty} b_n z^n, \tag{1.4}$$

the convolution (or Hadamard product) of

$$F(z) = -(\log \frac{f(z)}{z}) + (1 + 2\gamma_1)z, \tag{1.5}$$

and $G(z)$ denoted by $F * G$, is defined by

$$H(z) = F * G := z - \sum_{n=2}^{+\infty} 2\gamma_n b_n z^n. \tag{1.6}$$

We denote by $\Pi(\eta, \beta)$ and $Q(\eta, \beta)$ consisting of the functions $H(z) = F * G$ in N which satisfy

$$Re\{\frac{\frac{zH'(z)}{H(z)}}{\eta\frac{zH'(z)}{H(z)} + (1-\eta)}\} > \beta \tag{1.7}$$

and

$$Re\{\frac{1 + \frac{zH''(z)}{H'(z)}}{1 + \eta\frac{zH''(z)}{H'(z)}}\} > \beta\,, \quad 0 \le \beta < 1, 0 \le \eta < 1, \tag{1.8}$$

respectively. Also the functions $H(z)$ in N are said to be in the class $\Lambda(\eta, \beta, \psi)$, if there exists a function $\psi(z) \in N$ such that

$$Re\{\frac{\frac{zH'(z)}{\psi(z)}}{\eta\frac{zH'(z)}{\psi(z)} + (1-\eta)}\} > \beta. \tag{1.9}$$

For these subclassess we prove some interesting theorems include coefficient bounds, inclusion results, extreme points and property of convex sets.

Several other interesting subclasses of univalent functions were investigated recently, for example, by Ghanim and Darus (Ghanim & Darus, 2008), Prajapat and Goyal (Prajapat & Goyal, 2009), Acu and Owa (Acu & Owa, 2000) and etc. See also (Najafzadeh & Kulkarni, 2006) and (Najafzadeh & Ebadian, 2009).

2. Main result

Theorem 2.1. *If $H(z) \in \Lambda(\eta, \beta, \psi)$, then*

$$\sum_{n=2}^{+\infty} [2\gamma_n b_n(1 - \eta\beta) - \beta(1-\eta)c_n] \le 1 - \beta. \tag{2.1}$$

Proof. Since $H(z) \in \Lambda(\eta,\beta,\psi)$, then there exists a function $\psi(z) = z - \sum_{n=2}^{+\infty} c_n z^n \in N$ such that (1.9) holds true. By putting (1.6) and $H'(z) = (F * g)' = 1 - \sum_{n=2}^{+\infty} 2\gamma_n b_n z^n$ in (1.9) we get $Re\{\frac{1-\sum_{n=2}^{+\infty} 2\gamma_n b_n z^{n-1}}{1-\sum_{n=2}^{+\infty}(2\eta\gamma_n b_n+(1-\eta)c_n)z^{n-1}}\} > \beta$. By choosing the values of z on the real axis so that $\frac{z(F*G)'}{\psi(z)}$ is real and letting $r \to 1^-$ through real values, we have $\frac{1-\sum_{n=2}^{+\infty} 2\gamma_n b_n}{1-\sum_{n=2}^{+\infty}(2\eta\gamma_n b_n+(1-\eta)c_n)} \geq \beta$, or equivalently $\sum_{n=2}^{+\infty}[2\gamma_n b_n(1-\eta\beta) - \beta(1-\eta)c_n] \leq 1-\beta$. Now the proof is complete. □

Theorem 2.2. *If $H(z) \in Q(\eta,\beta)$, then $\sum_{n=2}^{+\infty} 2\gamma_n b_n(1+\eta(n-1)+\beta n^2) \leq 1-\beta$.*

Proof. Since $H(z) \in Q(\eta,\beta)$, then by (1.6) and (1.8) we get $Re\{\frac{1-\sum_{n=2}^{+\infty} 2n^2\gamma_n b_n z^{n-1}}{1-\sum_{n=2}^{+\infty} 2n\gamma_n(1+\eta(n-1))z^{n-1}}\} > \beta$. By choosing the values of z on the real axis so that $\frac{z(F*G)''}{(F*G)'}$ is real and letting $r \to 1^-$ though real values we have $\frac{1-\sum_{n=2}^{+\infty} 2n^2\gamma_n b_n}{1-\sum_{n=2}^{+\infty} 2n\gamma_n b_n(1+\eta(n-1))} > \beta$. The above inequality gives the required result. □

Definition 2.1. A function $H(z) \in N$ is said to be in $W(\eta,\beta)$, if there exists a function $\psi(z) = z - \sum_{n=2}^{+\infty} c_n z^n$ such that

(a) The condition (2.1) holds true;

(b) For every n, $2\gamma_n b_n - c_n \geq 0$.

In the next theorem we prove an inclusion property.

Theorem 2.3. *$W(\eta,\beta) \subseteq \Lambda(\eta,\beta,\psi)$.*

Proof. Let $H(z) \in W(\eta,\beta)$, we must show that $H(z) \in \Lambda(\eta,\beta,\psi)$ or equivalently the condition (1.9) holds. But

$$\left|\frac{\frac{z(F*G)'}{\psi(z)}}{\eta\frac{z(F*G)'}{\psi(z)}+(1-\eta)}-1\right| = \left|\frac{1-\sum_{n=2}^{+\infty} 2\gamma_n b_n z^{n-1}}{1-\sum_{n=2}^{+\infty}(2\eta\gamma_n b_n+(1-\eta)c_n)z^{n-1}}-1\right| = \left|\frac{(\eta-1)\sum_{n=2}^{+\infty}(2\gamma_n b_n - c_n)z^{n-1}}{1-\sum_{n=2}^{+\infty}(2\eta\gamma_n b_n+(1-\eta)c_n)z^{n-1}}\right|$$

$$\leq \frac{(1-\eta)\sum_{n=2}^{+\infty}(2\gamma_n b_n - c_n)}{1-\sum_{n=2}^{+\infty}(2\eta\gamma_n b_n+(1-\eta)c_n)}.$$

If (a) holds , above fraction is bounded above by $1-\alpha$ and hence (1.9) is satisfied. So $H(z) \in \Lambda(\eta,\beta,\psi)$. □

Remark. By putting $\psi(z) = G(z)$, in the last Theorem we obtain $\Pi(\eta,\beta) \subseteq W(\eta,\beta)$, and also by putting $\psi(z) = G(z)$ in (2.1) we have $\sum_{n=2}^{+\infty}[2\gamma_n(1-\eta\beta) - \beta(1-\eta)]b_n \leq 1-\beta$. This is the necessary and sufficient condition for functions $H(z) \in N$ to be in the class $\Pi(\eta,\beta)$.

3. Coefficient estimates and Distortion bounds for functions in $W(\eta,\beta)$

In this section we find coefficient bounds and verify distortion Theorem for the class $W(\eta,\beta)$.

Remark. If $H(z)$ be in the class $W(\eta,\beta)$, then

$$\sum_{n=2}^{+\infty}\gamma_n b_n \leq \frac{n(1-\beta)+\beta(1-\eta)}{2(1-\eta\beta)}. \tag{3.1}$$

Proof. From definition of $W(\eta,\beta)$ and taking $\psi(z)=z-\sum_{n=2}^{+\infty}c_n z^n$, we have $\sum_{n=2}^{+\infty}(1-\eta\beta)(2\gamma_n b_n)\leq 1-\beta+\beta(1-\eta)c_n$. If $c_n\leq\frac{1}{n}$ $(\forall n)$, thus we have $\sum_{n=2}^{+\infty}\gamma_n b_n\leq\frac{n(1-\beta)+\beta(1-\eta)}{2(1-\eta\beta)}$. □

Remark. The function $H_n(z)=z-\frac{n(1-\beta)+\beta(1-\eta)}{2(1-\eta\beta)}z^n$ is an extremal function for the class $W(\eta,\beta)$.

Theorem 3.1. *Let $H(z)=F*G$ be in the class $W(\eta,\beta)$, then for $|z|\leq r<1$*

$$r-\frac{2-\beta-\beta\eta}{4(1-\eta\beta)}r^2\leq|F*G|\leq r+\frac{2-\beta-\beta\eta}{4(1-\eta\beta)}r^2 \tag{3.2}$$

Proof. Since

$$H(z)=F*G=z-\sum_{n=2}^{+\infty}2\gamma_n b_n z^n, \tag{3.3}$$

so by (2.1) we get $\sum_{n=2}^{+\infty}2\gamma_n b_n(1-\eta\beta)-\beta(1-\eta)c_n\leq 1-\beta$. Since $c_n\leq\frac{1}{n}\leq\frac{1}{2}$ we have $\sum_{n=2}^{+\infty}2n\gamma_n b_n(1-\eta\beta)\leq\frac{\beta(1-\eta)}{2}+1-\beta$, or $2\sum_{n=2}^{+\infty}2n\gamma_n b_n(1-\eta\beta)\leq 2-\beta-\beta\eta$, or $2\sum_{n=2}^{+\infty}2\gamma_n b_n\leq 2\sum_{n=2}^{+\infty}n\gamma_n b_n\leq\frac{2-\beta-\beta\eta}{2(1-\eta\beta)}$, or $\sum_{n=2}^{+\infty}2\gamma_n b_n\leq\frac{2-\beta-\beta\eta}{4(1-\eta\beta)}$. From this inequality and (3.3) we have $|F*G|\leq|z|+\sum_{n=2}^{+\infty}2\gamma_n b_n|z|^n$ $\leq r+\frac{2-\beta-\beta\eta}{4(1-\eta\beta)}r^2$, and $|F*G|\geq r-\frac{2-\beta-\beta\eta}{4(1-\eta\beta)}r^2$. □

Theorem 3.2. *The class $W(\eta,\beta)$ is convex.*

Proof. Let $H_1(z)$ and $H_2(z)$ be in the class $W(\eta,\beta)$ with respect to functions $\psi_1(z)=z-\sum_{n=2}^{+\infty}c_n z^n$ and $\psi_2(z)=z-\sum_{n=2}^{+\infty}c'_n z^n$. For $0\leq j\leq 1$ we must show that $H(z)=jH_1(z)+(1-j)H_2(z)$ belongs to $W(\eta,\beta)$ with respect to $\psi(z)=j\psi_1(z)+(1-j)\psi_2(z)$. But $H_1(z)=z-\sum_{n=2}^{+\infty}2\gamma_n b_n z^n$, $H_2(z)=z-\sum_{n=2}^{+\infty}2\gamma_n b'_n z^n$, and $H(z)=z-\sum_{n=2}^{+\infty}s_n(j)z^n$, where $s_n(j)=2\gamma_n(jb_n+(1-j)b'_n)$. Also $\psi(z)=z-\sum_{n=2}^{+\infty}r_n(j)z^n$ where $r_n(j)=jc_n+(1-j)c'_n$.

The function $H(z)$ will belong to $W(\eta,\beta)$ if

(i) $\sum_{n=2}^{+\infty}[s_n(j)(1-\eta\beta)-\beta(1-\eta)r_n(j)]\leq 1-\beta$,

(ii) $s_n(j)-r_n(j)\geq 0$ for every n.

Since H_1 and H_2 are in $W(\eta,\beta)$ then $2\gamma_n b_n-c_n\geq 0$ and $2\gamma_n b'_n-c'_n\geq 0$, for all n. With direct calculation since $0\leq j\leq 1$ we have, $s_n(j)-r_n(j)=2\gamma_n(jb_n+(1-j)b'_n)-(jc_n+(1-j)c'_n)=j(2\gamma_n b_n-c_n)+(1-j)(2\gamma_n b'_n-c'_n)\geq 0$. Also $\sum_{n=2}^{+\infty}[s_n(j)(1-\eta\beta)-\beta(1-\eta)r_n(j)]=j\sum_{n=2}^{+\infty}2\gamma_n b_n(1-\eta\beta)-\beta(1-\eta)c_n+(1-j)\sum_{n=2}^{+\infty}2\gamma_n b'_n(1-\eta\beta)-\beta(1-\eta)c'_n\leq j(1-\beta)+(1-j)(1-\beta)=1-\beta$. Now the proof is complete. □

References

Acu, M. and S. Owa (2000). On some subclass of univalent functions. *Journal of Inequalities in Pure and Applied Mathematics* **6**(3), 1–14.

Ghanim, F. and M. Darus (2008). On new subclass of analytic univalent function with negative coefficient I. *Int. J. Contemp. Math. Sciences* **3**(27), 1317–1329.

Girela, D. (2000). Logarithmic coefficients of univalent functions. *Annals Acad. Sci. Fenn. Math. Series 1, Mathematica* **25**(2), 337–350.

Najafzadeh, Sh. and A. Ebadian (2009). Neighborhood and partial sum property for univalent holomorphic functions in terms of Komatu operator. *Acta Universitatis Apulensis* **25**(19), 81–90.

Najafzadeh, Sh. and S. R. Kulkarni (2006). Convex subclass of starlike functions in terms of combination of integral operators. *Int. Review of pure and appl. Math.* **2**(1), 25–34.

Prajapat, J. K. and S. P. Goyal (2009). Application of Srivastava-Attiya operator to the classes of strongly starlike and strongly convex functions. *J. Math. Ineq.* **3**(1), 129–137.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764, EISSN 2247-6202)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 90–98

On Ideal Convergent Difference Double Sequence Spaces in n-Normed Spaces Defined by Orlicz Function

Vakeel A. Khan[a,*], Sabiha Tabassum[a]

[a]*Department of Mathematics A.M.U. Aligarh-202002, India*

Abstract

The main aim of this paper is to define the generalized difference double sequence spaces ${}_2W^I(M, \|.,...,.\|, \Delta_m^n, p)$, ${}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p)$ and ${}_2W_\infty^I(M, \|.,...,.\|, \Delta_m^n, p)$ defined over a n-normed space $(X, \|.,...,.\|)$. Here we also study their properties and establish some inclusion relations.

Keywords: Double sequence spaces, n-norm, Orlicz Function, difference sequence spaces, I convergence.
2010 MSC: 46E30, 46E40, 46B20.

1. Introduction

The notion of ideal convergence was introduced first by Kostyrko et-al- (Kostyrko *et al.*, 2000) as an interesting generalization of statistical convergence (Khan & Tabassum, 2012) which was further studied in topological spaces. A family $I \subset 2^Y$ of subsets of a nonempty set Y is said to be an ideal in Y if

1. $\emptyset \in I$;
2. $A, B \in I$ imply $A \cup B \in I$;
3. $A \in I, B \subset A$ imply $B \in I$,

while an admissible ideal I further satisfies $\{x\} \in I$ for each $x \in Y$ (Kostyrko *et al.*, 2000, 2005; Savas, 2010).
Given $I \subset 2^{\mathbb{N}}$ be a nontrivial ideal in $\mathbb{N}$. Let X be a normed space. The sequence (x_j) in X is said to be $I-$*convergent* to $\xi \in X$, if for each $\varepsilon > 0$ the set $A(\varepsilon) = \{j \in \mathbb{N} : \|x_j - \xi\| \geq \varepsilon\}$ belongs to I (Khan & Tabassum, 2010).

*Corresponding author
Email addresses: vakhanmaths@gmail.com (Vakeel A. Khan), sabihatabassum@math.com (Sabiha Tabassum)

The concept of 2-normed spaces was initially introduced by G̈ahler (Gähler, 1963) in the mid of 1960's as an interesting nonlinear generalization of a normed linear space. Since then, many researchers have studied this concept and obtained various results, see for instance (Gunawan & Mashadi, 2001; Khan & Tabassum, 2010; Savas, 2010).

Recall (Khan & Tabassum, 2012) that an *Orlicz Function* is a function $M : [0, \infty) \to [0, \infty)$ which is continuous, nondecreasing and convex with $M(0) = 0$, $M(x) > 0$ for $x > 0$ and $M(x) \to \infty$, as $x \to \infty$. If convexity of M is replaced by $M(x+y) \leq M(x)+M(y)$, then it is called a *Modulus funtion* (Maddox, 1986).

Let w be the space of all sequences. Lindenstrauss and Tzafriri (Lindenstrauss & Tzafiri, 1971) used the idea of Orlicz sequence space. Let

$$l_M := \left\{x \in w : \sum_{k=1}^{\infty} M\left(\frac{|x_k|}{\rho}\right) < \infty, \text{ for some } \rho > 0\right\}$$

is Banach space with respect to the norm

$$\|x\|_M := \inf\left\{\rho > 0 : \sum_{k=1}^{\infty} M\left(\frac{|x_k|}{\rho}\right) \leq 1\right\}.$$

Orlicz function has been studied by V. A. Khan (Khan, 2008*a*,*b*) and many others.

Let $n \in \mathbb{N}$ and X be a real vector space of dimension d, where $n \leq d$. An n-norm on X is a function $\|., ..., .\| : X \times X \times ... \times X \to \mathbb{R}$ which satisfies the following four conditions:

1. $\|x_1, x_2, ..., x_n\| = 0$ if and only if $x_1, x_2, ..., x_n$ are linearly dependent,
2. $\|x_1, x_2, ..., x_n\|$ is invariant under permutation,
3. $\|\alpha x_1, x_2, ..., x_n\| = |\alpha|\|x_1, x_2, ..., x_n\|$, for any $\alpha \in \mathbb{R}$,
4. $\|x + x', x_2, ..., x_n\| \leq \|x, x_2, ..., x_n\| + \|x', x_2, ..., x_n\|$.

The pair $(X, \|., ., ..., .\|)$ is called an n-normed space (Savas, 2011).

Example 1.1. (see (Savas, 2011)). As a standard example of a n-normed space we may take R^n being equipped with the n-norm $\|x_1, x_2, ..., x_n\|_E$ = the volume of the n-dimensional parallelopiped spaned by the vectors $x_1, x_2, ..., x_{n-1}, x_n$ which may be given explicitly by the formula

$$\|x_1, x_2, ..., x_n\|_E = \begin{vmatrix} \langle x_1, x_2\rangle & ... & \langle x_1, x_n.\rangle \\ . & ... & \\ . & ... & \\ \langle x_n, x_1\rangle & ... & \langle x_n, x_n\rangle \end{vmatrix}.$$

where $\langle ., .\rangle$ denotes inner product.

Example 1.2. (see (Savas, 2011)). Let $(X, \|., ., ..., .\|)$ be an n-normed space of dimension $d \geq n \geq 2$ and $\{a_1, a_2, ..., a_n\}$ be a linearly independent set in X. Then the following function $\|., ., ..., .\|_\infty$ defined by

$$\|x_1, x_2, ..., x_{n-1}, x_n\|_\infty = \max\{\|x_1, x_2, ..., x_{n-1}, a_i\| : i = 1, 2, ..., n\}$$

defines an $(n-1)$−norm on X with respect to $\{a_1, a_2, ..., a_n\}$.

Definition 1.1. (see (Savas, 2011)). A sequence (x_j) in an n-normed space $(X, \|., ., ..., .\|)$ is said to be converge to some $L \in X$ in the n-norm if

$$\lim_{j\to\infty} \|x_j - L, x_1, ..., x_{n-1}\| = 0, \quad \text{for every } x_1, ..., x_{n-1} \in X.$$

Example 1.3. (see (Khan & Tabassum, 2010)). A sequence (x_j) in an n-normed space $(X, \|., ., ..., .\|)$ is said to be Cauchy with respect to the n-norm if

$$\lim_{j,k\to\infty} \|x_j - x_k, x_1, ..., x_{n-1}\| = 0, \quad \text{for every } x_1, ..., x_{n-1} \in X.$$

If every Cauchy sequence in X converges to some $L \in X$, then X is said to be complete with respect to the n-norm. Any complete n-normed space is said to be n-Banach space.

Let w, l_∞, c and c_0 denote the spaces of all, bounded, convergent and null sequences $x = (x_k)$ with complex terms, respectively, normed by

$$\|x\| = \sup_k |x_k|.$$

Kizmaz (Kizmaz, 1981), defined the difference sequences $l_\infty(\Delta), c(\Delta)$ and $c_0(\Delta)$ as follows:

$$Z(\Delta) = \{x = (x_k) : (\Delta x_k) \in Z\},$$

for $Z = l_\infty, c$ and c_0, where $\Delta x = (\Delta x_k) = (x_k - x_{k+1})$, for all $k \in \mathbb{N}$.

The above spaces are Banach spaces, normed by

$$\|x\|_\Delta = |x_1| + \sup_k |\Delta x_k|.$$

The notion of difference sequence spaces was generalized by Et. and Colak (Et & Colak, 1995) as follows:

$$Z(\Delta^n) = \{x = (x_k) : (\Delta^n x_k) \in Z\},$$

for $Z = l_\infty, c$ and c_0, where $n \in \mathbb{N}, (\Delta^n x_k) = (\Delta^{n-1} x_k - \Delta^{n-1} x_{k+1})$ and so that

$$\Delta^n x_k = \sum_{v=0}^{n} (-1)^v \binom{n}{v} x_{k+v}.$$

In 2005, Tripathy and Esi (Tripathy & Esi, 2006), introduced the following new type of difference sequence spaces:

$$Z(\Delta_m) = \{x = (x_k) \in w : \Delta_m x \in Z\}, \text{ for } Z = l_\infty, c \text{ and } c_0$$

where $\Delta_m x = (\Delta_m x_k) = (x_k - x_{k+m})$, for all $k \in \mathbb{N}$.

Later on, Tripathy, Esi and Tripathy (B. C. Tripathy & Tripathy, 2005), generalized the above notions and unified these as follows:
Let m, n be non negative integers, then for Z a given sequence space we have

$$Z(\Delta_m^n) = \{x = (x_k) \in w : (\Delta_m^n x_k) \in Z\}$$

where $\Delta_m^n x = (\Delta_m^n x_k) = (\Delta_m^{n-1} x_k - \Delta_m^{n-1} x_{k+m})$ and $\Delta_m^\circ x_k = x_k$ for all $k \in \mathbb{N}$. The difference operator is equivalent to the binomial representation

$$\Delta_m^n x_k = \sum_{v=0}^{n} (-1)^v \binom{n}{v} x_{k+mv}.$$

A *paranorm* is a function $g : X \to \mathbb{R}$ which satisfies the following axioms:
For any $x, y, x_0 \in X$, $\lambda, \lambda_0 \in \mathbb{C}$:

(i) $g(\theta) = 0$;
(ii) $g(x) = g(-x)$;
(iii) $g(x + y) \leq g(x) + g(y)$
(iv) the scalar multiplication is continuous, that is $\lambda \to \lambda_0$, $x \to x_0$ imply $\lambda x \to \lambda_0 x_0$.

Throughout, a double sequence $x = (x_{jk})$ is a double infinite array of elements x_{jk}. for $j, k \in \mathbb{N}$. Double sequences have been studied by V. A. Khan and S. Tabassum (Khan & Tabassum, 2012; V. & Tabassum, 2011; Khan & Tabassum, 2011, 2010), Moricz and Rhoades (Moricz & Rhoades, 1952) and many others.

Definition 1.2. (see (Khan & Tabassum, 2010)). A double sequence space X is said to be *Solid (Normal)*, if $(\alpha_{jk} x_{jk}) \in X$ whenever $(x_{jk}) \in X$ and for all double sequence (α_{jk}) of scalars with $|\alpha_{jk}| \leq 1$ for all $j, k \in \mathbb{N}$.

2. Main Results

In 2010 E. Savas (Savas, 2010) introduced certain new sequence spaces using ideal convergence in 2-normed spaces. Later on V. A. Khan and S. Tabassum (Khan & Tabassum, 2010) introduced similar kind of double sequence spaces using difference operator in 2-normed spaces. In this paper we generalized these sequence spaces in n-normed spaces.

Let $p = (p_{jk})$ be any bounded sequence of positive numbers, m, n be non-negative integers and let I be an admissible ideal of $\mathbb{N}$. Let ${}_2W(n - X)$ be the space of X-valued double sequence spaces defined over a n-normed space $(X, \|., ..., .\|)$. Then for an Orlicz function M we define the following sequence spaces:

$${}_2W^I(M, \|., ..., .\|, \Delta_m^n, p) = \Big\{x = (x_{jk}) \in {}_2W(n - X) : \forall \varepsilon > 0 \text{ the set } \Big\{(j, k) \in N \times N :$$

$$\lim_{j,k \to \infty} \left(M\left(\left\|\frac{\Delta_m^n x_{jk} - L}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \varepsilon\Big\} \in I, \text{ for some } \rho > 0, L \in X, z_1, z_2, ..., z_{n-1} \in X\Big\}.$$

$${}_2W_0^I(M, \|., ..., .\|, \Delta_m^n, p) = \Big\{x = (x_{jk}) \in {}_2W(n - X) : \forall \varepsilon > 0 \text{ the set}\Big\{(j, k) \in N \times N :$$

$$\lim_{j,k\to\infty}\left(M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \varepsilon\Big\} \in I, \text{ for some } \rho > 0, z_1, z_2, ..., z_{n-1} \in X\Big\}.$$

$${}_2W_\infty^I(M, \|., ..., .\|, \Delta_m^n, p) = \Big\{x = (x_{jk}) \in {}_2W(n-X) : \exists K > 0 \text{ s.t. } \Big\{(j,k) \in N \times N :$$

$$\sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq K\Big\} \in I, \text{ for some } \rho > 0, z_1, z_2, ..., z_{n-1} \in X\Big\}$$

where

$$(\Delta_m^n x_{jk}) = (\Delta_m^{n-1} x_{jk} - \Delta_m^{n-1} x_{j+1,k} - \Delta_m^{n-1} x_{j,k+1} + \Delta_m^{n-1} x_{j+1,k+1})$$

and

$$(\Delta_m^0 x_{jk}) = x_{jk} \quad \text{for all } \ j,k \in \mathbb{N},$$

which is equivalent to the following binomial representation:

$$\Delta_m^n x_{jk} = \sum_{u=0}^{n}\sum_{v=0}^{n}(-1)^{u+v}\binom{n}{u}\binom{n}{v}x_{j+mu,k+mv}.$$

and $\Delta x_{j,k} = x_{j,k} - x_{j+1,k} - x_{j,k+1} + x_{j+1,k+1}$.
The following inequality will be used throughout the paper. Let $p_{j,k}$ be a double sequence of positive real numbers with $0 < p_{jk} \leq \sup_{j,k} p_{jk} = H$, and let $D = \max\{1, 2^{H-1}\}$. Then for the factorable sequences (a_{jk}) and (b_{jk}) in the complex plane, we have

$$|a_{jk} + b_{jk}|^{q_{jk}} \leq D(|a_{jk}|^{q_{jk}} + |b_{jk}|^{q_{jk}})$$

Theorem 2.1. *If* $\{\Delta_m^n x_{jk}, z_1, z_2, ..., z_{n-1}\}$ *is a linearly independent set in* $(X.\|., ..., .\|)$ *for all but finite* j,k *where* $x = (x_{jk}) \in {}_2W(n-X)$ *and* $\inf_{j,k} p_{jk} > 0$, *then*

(i) $\lim_{j,k\to\infty}\left[M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} = 0$, for every $\rho > 0$,

(ii) $\lim_{j,k\to\infty}\left[M\left(\left\|\frac{\Delta_m^n x_{jk}-L}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} < \infty$, for every $\rho > 0$.

Proof. **(i).** Assume that $\{\Delta_m^n x_{jk}, z_1, z_2, ..., z_{n-1}\}$ is a linearly independent set in $(X.\|., ..., .\|)$ for all but finite j,k. Then we have $\|\Delta_m^n x_{jk}, z_1, z_2, ..., z_{n-1}\| \to 0$ as $j,k \to \infty$.
Since M is continuous and $0 < p_{jk} \leq \sup p_{jk} < \infty$, for each j,k, we have

$$\lim_{j,k\to\infty}\left[M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} = 0, \text{ for every } \rho > 0.$$

(ii). Proof of this part is similar to part **(i)**. □

Theorem 2.2. ${}_2W^I(M, \|., ..., .\|, \Delta_m^n, p)$, ${}_2W_0^I(M, \|., ..., .\|, \Delta_m^n, p)$ *and* ${}_2W_\infty^I(M, \|., ..., .\|, \Delta_m^n, p)$ *are linear spaces.*

Proof. We prove the assertion for ${}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p)$ the others can be proved similarly. Assume that $x = (x_{jk})$ and $y = (y_{jk}) \in {}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p)$ and $\alpha, \beta \in \mathbb{R}$, so

$$\left\{(j,k) \in N \times N : \lim_{j,k\to\infty} \left(M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho_1}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \varepsilon\right\} \in I, \text{ for some } \rho_1 > 0, \tag{2.1}$$

$$\left\{(j,k) \in N \times N : \lim_{j,k\to\infty} \left(M\left(\left\|\frac{\Delta_m^n y_{jk}}{\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \varepsilon\right\} \in I, \text{ for some } \rho_2 > 0, \tag{2.2}$$

Since $\|.,...,.\|$ is a n-norm, and M is an Orlicz function the following inequality holds:

$$\begin{aligned}
&\lim_{j,k\to\infty} \left(M\left(\left\|\tfrac{\Delta_m^n(\alpha x_{jk}+\beta y_{jk})}{|\alpha|\rho_1+|\beta|\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \\
&\leq D \lim_{j,k\to\infty} \left[\tfrac{|\alpha|\rho_1}{|\alpha|\rho_1+|\beta|\rho_2} M\left(\left\|\tfrac{\Delta_m^n x_{jk}}{\rho_1}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} \\
&+D \lim_{j,k\to\infty} \left[\tfrac{|\beta|\rho_2}{|\alpha|\rho_1+|\beta|\rho_2} M\left(\left\|\tfrac{\Delta_m^n y_{jk}}{\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} \\
&\leq DF \lim_{j,k\to\infty} \left[M\left(\left\|\tfrac{\Delta_m^n x_{jk}}{\rho_1}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} \\
&+DF \lim_{j,k\to\infty} \left[M\left(\left\|\tfrac{\Delta_m^n y_{jk}}{\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}}
\end{aligned} \tag{2.3}$$

where

$$F = \max\left[1, \left(\frac{|\alpha|}{\alpha\rho_1 + |\beta|\rho_2}\right)^H, \left(\frac{|\beta|}{\alpha\rho_1 + |\beta|\rho_2}\right)^H\right] \tag{2.4}$$

From the above inequality, we get

$$\begin{aligned}
&\left\{(j,k) \in N \times N : \lim_{j,k\to\infty} \left(M\left(\left\|\tfrac{\Delta_m^n \alpha x_{jk}+\Delta_m^n \beta y_{jk}}{|\alpha|\rho_1+|\beta|\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \varepsilon\right\} \\
&\subseteq \left\{(j,k) \in N \times N : DF \lim_{j,k\to\infty} \left(M\left(\left\|\tfrac{\Delta_m^n x_{jk}}{\rho_1}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \tfrac{\varepsilon}{2}\right\} \\
&\cup\left\{(j,k) \in N \times N : DF \lim_{j,k\to\infty} \left(M\left(\left\|\tfrac{\Delta_m^n y_{jk}}{\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \tfrac{\varepsilon}{2}\right\}.
\end{aligned} \tag{2.5}$$

The sets on the right hand side belong to I and this completes the proof. □

Theorem 2.3. *For any fixed $(j,k) \in N \times N$, ${}_2W_\infty^I(M, \|.,...,.\|, \Delta_m^n, p)$ is paranomed space with respect to the paranorm defined by:*

$$g(x) = \inf_{j,k}\left\{\rho^{\frac{p_{jk}}{H}} : \left(\sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}}\right)^{\frac{1}{H}} \leq 1, \forall z_1, z_2, ..., z_{n-1} \in X\right\}. \tag{2.6}$$

Proof. (**i**) $x = \theta$ implies that then $\|0, z_1, z_2..., z_{n-1}\| = 0$ since the set containing 0 is linearly dependent. Also $M(0) = 0$ implies that $g(\theta) = 0$.

(ii) $g(x) = g(-x)$

(iii) Let $x = (x_{jk}), y = (y_{jk}) \in {}_2W_\infty^I(M, \|., ..., .\|, \Delta_m^n, p)$.

Then there exists $\rho_1, \rho_2 > 0$ such that: $\sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho_1}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \leq 1$ and

$$\sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n y_{jk}}{\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \leq 1 \tag{2.7}$$

for each $z_1, z_2, ..., z_{n-1} \in X$.

Let $\rho = \rho_1 + \rho_2$. Then by convexity of Orlicz function we have:

$$\begin{aligned}\sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n x_{jk}+\Delta_m^n y_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right) &\leq \left(\frac{\rho_1}{\rho_1+\rho_2}\right)\sup_{j,k\geq 1} M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho_1}, z_1, z_2, ..., z_{n-1}\right\|\right)\\ &+\left(\frac{\rho_2}{\rho_1+\rho_2}\right)\sup_{j,k\geq 1} M\left(\left\|\frac{\Delta_m^n y_{jk}}{\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right).\end{aligned} \tag{2.8}$$

Thus $\sup_{j,k\geq 1} M\left(\left\|\frac{\Delta_m^n x_{jk}+\Delta_m^n y_{jk}}{\rho_1+\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)^{p_{jk}} \leq 1$ and hence

$$\begin{aligned}g(x+y) \leq \inf_{j,k}\left\{\rho_1^{\frac{p_{jk}}{H}} : \sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho_1}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \leq 1\right\}\\ + \inf_{j,k}\left\{\rho_2^{\frac{p_{jk}}{H}} : \sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n y_{jk}}{\rho_2}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \leq 1\right\}.\end{aligned} \tag{2.9}$$

The arbitrary ρ_1 and ρ_2 implies that $g(x + y) \leq g(x) + g(y)$.

(iv) Let $\alpha \to 0$ and $g(x^n - x) \to 0 \quad (n \to \infty)$

$$g(\alpha x) = \inf\left\{\left(\frac{\rho}{|\alpha|}\right)^{\frac{p_{jk}}{H}} : \sup_{j,k\geq 1}\left(M\left(\left\|\frac{\Delta_m^n \alpha x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \leq 1\right\}. \tag{2.10}$$

□

Theorem 2.4. *Let M, M_1, M_2, be Orlicz functions. Then we have*

(i) ${}_2W_0^I(M_1, \|., ..., .\|, \Delta_m^n, p) \subseteq {}_2W_0^I(M \circ M_1, \|., ..., .\|, \Delta_m^n, p)$
provided (p_{jk}) is such that $H_0 = \inf p_{jk} > 0$.

(ii) ${}_2W_0^I(M_1, \|., ..., .\|, \Delta_m^n, p) \cap {}_2W_0^I(M_2, \|., ..., .\|, \Delta_m^n, p) \subseteq {}_2W_0^I(M_1 + M_2, \|., ..., .\|, \Delta_m^n, p)$.

Proof. **(i).** For given $\varepsilon > 0$, first choose $\varepsilon_0 > 0$ such that $\max\{\varepsilon_0^H, \varepsilon_0^{H_0}\} < \varepsilon$. Now using the continuity of M choose $0 < \delta < 1$ such that $0 < t < \delta$, implies that $M(t) < \varepsilon_0$, Let $(x_{jk}) \in {}_2W_0^I(M_1, \|., .\|, \Delta_m^n, p)$. Now by definition:

$$A(\delta) = \left\{(j,k) \in N \times N : \lim_{j,k\to\infty}\left(M_1\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \geq \delta^H\right\} \in I. \tag{2.11}$$

Thus if $(j,k) \notin A(\delta)$ then

$$\left(M_1\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} \leq \delta^H, \qquad \forall j,k \in \mathbb{N}. \tag{2.12}$$

That is

$$\left(M_1\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} < \delta, \qquad \forall j,k \in \mathbb{N}. \tag{2.13}$$

Hence from above using continuity of M we must have

$$M\left(M_1\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)^{p_{jk}} < \varepsilon_0, \qquad \forall j,k \in \mathbb{N} \tag{2.14}$$

Which consequently implies that

$$\lim_{j,k\to\infty}\left[M\left(M_1\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)\right]^{p_{jk}} < \max\{\varepsilon_0^H, \varepsilon_0^{H_0}\} < \varepsilon. \tag{2.15}$$

This shows that

$$\left\{(j,k) \in N \times N : \lim_{j,k\to\infty}\left[M\left(M_1\left(\left\|\frac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right)\right]^{p_{jk}} \geq \varepsilon\right\} \subset A(\delta) \tag{2.16}$$

and so belongs to I. This completes the result.
(ii). Let $x_{jk} \in {}_2W_0^I(M_1, \|.,...,.\|, \Delta_m^n, p) \cap {}_2W_0^I(M_2, \|.,...,.\|, \Delta_m^n, p)$
Then the fact that

$$\begin{aligned}&\lim_{j,k\to\infty}\left[(M_1 + M_2)\left(\left\|\tfrac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} \leq D \lim_{j,k\to\infty}\left[M_1\left(\left\|\tfrac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} \\ &+D \lim_{j,k\to\infty}\left[M_2\left(\left\|\tfrac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}}.\end{aligned} \tag{2.17}$$

This gives the result. □

Theorem 2.5. *The sequence space* ${}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p), {}_2W_\infty^I(M, \|.,...,.\|, \Delta_m^n, p)$ *are Solid.*

Proof. We give the proof for ${}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p)$ only.

Let $(x_{jk}) \in {}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p)$ and let (α_{jk}) be a double sequence of scalars such that $|\alpha_{jk}| \leq 1$ for all $j,k \in \mathbb{N}$. Then we have

$$\begin{aligned}&\left\{(j,k) \in N \times N : \lim_{j,k\to\infty}\left[M\left(\left\|\tfrac{\Delta_m^n(\alpha_{jk}x_{jk})}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} \geq \varepsilon\right\} \\ &\subseteq \left\{(j,k) \in N \times N : E \lim_{j,k\to\infty}\left[M\left(\left\|\tfrac{\Delta_m^n x_{jk}}{\rho}, z_1, z_2, ..., z_{n-1}\right\|\right)\right]^{p_{jk}} \geq \varepsilon\right\} \in I.\end{aligned} \tag{2.18}$$

Where $E = \max_{j,k}\{1, |\alpha_{jk}|^H\}$. Hence $(\alpha_{jk}x_{jk}) \in {}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p)$ for all double sequence of scalars (α_{jk}) with $|\alpha_{jk}| \leq 1$ for all $j,k \in \mathbb{N}$ whenever $(x_{jk}) \in {}_2W_0^I(M, \|.,...,.\|, \Delta_m^n, p)$. □

Acknowledgments. The authors would like to record their gratitude to the reviewer for his careful reading and making some useful corrections which improved the presentation of the paper.

References

B. C. Tripathy, A. Esi and B. K. Tripathy (2005). A new type of generalized difference cesáro sequence spaces. *Soochow J. of Math.* **31**(03), 333–340.

Et, M. and R. Colak (1995). On generalized difference sequence spaces. *Soochow Jour. Math.* **21**(4), 377–386.

Gähler, S. (1963). 2-merische räme und ihre topological struktur. *Math. Nachr.* **28**(4), 115–148.

Gunawan, H. and Mashadi (2001). On finite dimensional 2-normed spaces. *Soochow Jour. Math.,* **27**(3), 631–639.

Khan, V. A. (2008*a*). On a new sequence space defined by Orlicz functions. *Commun. Fac. Sci. Univ. Ank. Series Al* **57**(2), 25–33.

Khan, V. A. (2008*b*). On a new sequence space related to the Orlicz sequence space. *J. Mathematics and its applications* **30**, 61–69.

Khan, V. A. and S. Tabassum (2010). On ideal convergent difference double sequence spaces in 2-normed spaces defined by Orlicz function. *JMI International Journal of Mathematical Sciences* **1**(2), 26–34.

Khan, V. A. and S. Tabassum (2011). Statistically convergent double sequence spaces in 2-normed spaces defined by Orlicz function. *Applied Mathematics* **2**(4), 398–402.

Khan, V. A. and S. Tabassum (2012). Statistically Pre-Cauchy double sequences and Orlicz functions. *Southeast Asian Bull.Math.* **36**(2), 61–69.

Kizmaz, H. (1981). On certain sequence spaces. *Canad. Math. Bull.*

Kostyrko, P., M. Ňacaj, T. Šalát and M.Sleziak (2005). *i*-convergence and extremal *i*-limit points. *Math. Slovaco*.

Kostyrko, P., T. Šalát and W. Wilczynski (2000). Lectures on analysis. *Canad. Math. Bull.*

Lindenstrauss, J. and L. Tzafiri (1971). On Orlicz sequence spaces. *Math. Slovaco* **10**, 379– 390.

Maddox, I. J (1986). Sequence spaces defined by modulus. *Math. Proc. Camb. Soc.* **100**(2), 161–166.

Moricz, F. and B. E. Rhoades (1952). Almost convergence of double sequences and strong regularity of summability matrices. *Math.Proc.Camb.Phil.Soc.* **104**, 283–294.

Savas, E. (2010). On some new sequence spaces in 2-normed spaces using ideal convergence and an Orlicz function. *Journal of Inequalities and Applications* **104**, 283–294.

Savas, E. (2011). Some new double sequence spaces defined by Orlicz function in n-normed space. *Journal of Inequalities and Applications* **104**, 283–294.

Tripathy, B. C. and A. Esi (2006). A new type of difference sequence spaces. *International Journal of Science and Technology* **01**(01), 11–14.

V., A., Khan and S. Tabassum (2011). On some new quasi almost δ^m–lacunary strongly p-convergent double sequences defined by Orlicz functions. *Journal of Mathematics and ApplicationsSoutheast Asian Bull.Math.* **34**, 45–52.

Theory and Applications of Mathematics & Computer Science

(ISSN 2067-2764, EISSN 2247-6202)
http://www.uav.ro/applications/se/journal/index.php/tamcs

Theory and Applications of Mathematics & Computer Science 3 (1) (2013) 99–106

Satellite Constellation Reconfiguration Using the Approximating Sequence Riccati Equations

Ashraf H. Owis[a]

[a]*Department of Astronomy, Space and Meteorology Cairo University*

Abstract

In this work we study the reconfiguration of a constellation of satellite. In this work we consider the non-linear feedback optimal control of the motion of a spacecraft under the influence of the gravitational attraction of a central body, the Earth in our case, and we would like to transfer the spacecraft from lower circular orbit to a higher one. Both orbits around the Earth are assumed to be circular and coplanar. We use both radial and tangential thrust control. The nonlinear dynamics of the system will be factorized in such a way that the new factorized system is accessible. The problem is tackled using the Approximating Sequence Riccati Equations (ASRE) method. The technique is based on Linear Quadratic Regulator (LQR) with fixed terminal state, which guarantees closed loop solution. The method is tested through GNSS circular constellation.

Keywords: Nonlinear feedback, linear quadratic regulator, approximation sequence Riccati equation, GNSS satellite.
2010 MSC: 49.

1. Introduction

In some instances, it is desirable to deploy a constellation in stages to gradually expand its capacity. This requires launching additional satellites and reconfiguring the existing on-orbit satellites (de Weck *et al.*, 2008). Also, a constellation might be re-structured and reconfigured after it is initially set for operational reasons.

The most common way of raising or lowering the orbit of a spacecraft is the low thrust orbit rendezvous approach, which is a nonlinear optimal control problem. Historically, there are several method to solve the nonlinear optimal control problem in both open and closed loop contexts. In the open loop context the problem can be solved via indirect and then direct method. The indirect method was developed through Pontryagin Maximum Principle (PMP) (Bryson & Ho, 1975),

*Corresponding author
Email address: aowis@eun.eg (Ashraf H. Owis)

(Pontryagin *et al.*, 1952). The direct method was developed using the Karush-Kuhn-Tucker(KKT) algebric equation (Enright & Conway, 1992).

One of the most common methods for solving the nonlinear optimal control problem in the closed loop context is the State Dependent Riccati Equations (SDRE) (Cimen, 2006), (Owis, 2013). The Approximating Sequence of Riccati Equations (ASRE) (Cimen, 2004) technique is an iterative approach to solve the nonlinear optimal control problem. The ASRE is developed (Topputo & Bernelli-Zazzera, 2012) using the state transition matrix. The guidance designed with these methods is obtained in an open-loop context. In other words, the optimal path, even if minimizing the prescribed performance index, is not able to respond to any perturbation that could alter the state of the spacecraft. Furthermore, if the initial conditions are slightly varied (e.g. the launch date changes), the optimal solution needs to be recomputed again. The outcome of the classical problem is in fact a guidance law expressed as a function of the time, the initial and final time, and u the control vector, respectively. We develop a closed loop approach. With this approach the solutions that minimize the performance index are also functions of the generic initial state x_0; the outcome is in fact a guidance law written as $u = u(x_0, t_0, t)$, $t \in [t_0, t_f]$. This represents a closed-loop solution: given the initial conditions (t_0, x_0) it is possible to extract the optimal control law that solves the optimal control problem. Moreover, if for any reason the state is perturbed and assumes the new value $(t_0^{'}, x_0^{'}) = (x_0 + \delta x, t_0 + \delta t)$, we are able to compute the new optimal solution by simply evaluating so avoiding the solution of another optimal control problem. This property holds by virtue of the closed loop characteristics of the control law that can be viewed as a one-parameter family of solutions. Due to such property, a trajectory designed in this way has the property to respond to perturbations acting during the transfer that continuously alter the state of the spacecraft. The optimal feedback control for linear systems with quadratic objective functions is addressed through the matrix Riccati equation: this is a matrix differential equation that can be integrated backward in time to yield the initial value of the Lagrange multipliers (Bryson & Ho, 1975). Recently, the nonlinear feedback control of circular coplanar low-thrust orbital transfers has been faced using continuous orbital elements feedback and Lyapunov functions (Chang & Marsden, 2002) and proved optimal by (Alizadah & Villac, 2011). Later on the problem has been solved using the primer vector approximation method (Haung, 2012).

The analytical low-thrust optimal feedback control problem is solved, with modulated inverse-square-distance, in the frame of a nonlinear vector field, the two-body dynamics, supported by a nonlinear objective function by applying a globally diffeomorphic linearizing transformation that rearranges the original problem into a linear system of ordinary differential equations and a quadratic objective function written in a new set of variables with radial thrust (Topputo *et al.*, 2008). In this work we consider the nonlinear feedback optimal control of the motion of a spacecraft under the influence of the gravitational attraction of a central body, the Earth in our case, and we would like to transfer the spacecraft from lower to higher orbit. Both lower and higher orbits around the Earth are assumed to be circular and coplanar. We use both radial and tangential thrust control. The nonlinear dynamics of the system will be factorized in such a way that the new factorized system is accessible. The problem is tackled using the Approximating Sequence Riccati Equation (ASRE) method. The technique is based on Linear Quadratic Regulator (LQR) with fixed terminal state. The method is applied to GNSS circular constellation Figure 1.

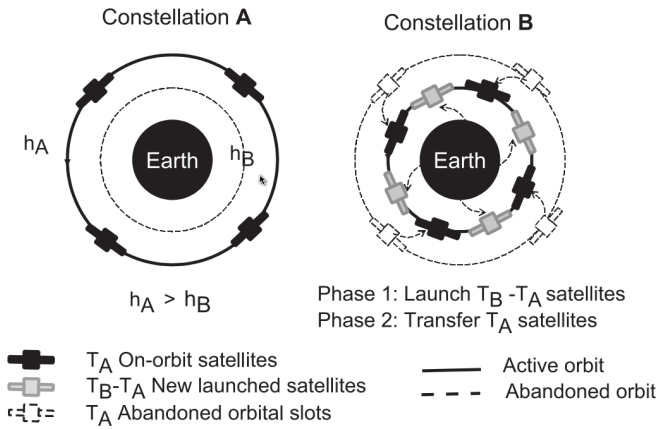


Figure 1. Constellation reconfiguration.

Linear Quadratic Regulator(LQR) with Fixed Terminal State.

Consider the following system with linear dynamics and quadratic performance index as follows:

$$\dot{X} = AX + BU, \qquad X(t_0) = X_0 \in \mathbb{R}^n, \tag{1.1}$$

the following performance index

$$J = X_f^T Q_f X_f + \frac{1}{2}\int_{t_0}^{t_f} [X^T QX + U^T RU]dt, \tag{1.2}$$

where A , B , Q , and R are constant coefficients matrices of the suitable dimensions. we have to find the m-dimensional control functions $U(t)$, $t \in [t_0\ t_f]$ which minimizes the J, which is an open loop (with t_0 fixed) optimal control. We optimize the performance index J, by adjoining the dynamics and the performance index (integrand) to form the Hamiltonian:

$$H(X, \lambda, U, t) = \frac{1}{2}(X^T QX + U^T RU) + \lambda^T (A(t)X + B(t)U),$$

where the Lagrange multiplier λ is called the adjoint variable or the costate. The necessary conditions for optimality are:

1. $\dot{X} = H_\lambda = A(t)X + B(t)U, \qquad X(t_0) = X_0,$
2. $\dot{\lambda} = -H_x = -QX - A^T\lambda, \qquad \lambda(t_f) = Q_f X_f,$
3. $H_u = 0 \Longrightarrow RU + B^T\lambda = 0 \Longrightarrow U^\star = -R^{-1}B^T\lambda.$

To find the minimum solution we have to check for $H_{uu} = \dfrac{\partial^2 H}{\partial \lambda^2} > 0$ or equivalently $R > 0$. Now we have that $\dot{X} = AX + BU^\star = AX - BR^{-1}B^T\lambda$, which can be combined to the the equation of the costate as follows

$$\begin{bmatrix} \dot{X} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} X \\ \lambda \end{bmatrix}, \tag{1.3}$$

which is called the Hamiltonian matrix, it represents a $2n$ boundary value problem with $X(t_0) = X_0$ and, $\lambda(t_f) = Q_f X_f$.

We can solve this $2n$ boundary value problem using the transition matrix method as follows. Let's define a transition matrix

$$\phi_(t_1, t_0) = \begin{bmatrix} \phi_{11}(t_1, t_0) & \phi_{12}(t_1, t_0) \\ \phi_{21}(t_1, t_0) & \phi_{22}(t_1, t_0) \end{bmatrix},$$

we use this matrix to relate the current values of X and λ to the final values X_f and λ_f as follows

$$\begin{bmatrix} X \\ \lambda \end{bmatrix} = \begin{bmatrix} \phi_{11}(t, t_f) & \phi_{12}(t, t_f) \\ \phi_{21}(t, t_f) & \phi_{22}(t, t_f) \end{bmatrix} \begin{bmatrix} X(t_f) \\ \lambda(t_f) \end{bmatrix},$$

so we have $X = \phi_{11}(t, t_f)X(t_f) + \phi_{12}(t, t_f)\lambda(t_f) = [\phi_{11}(t, t_f) + \phi_{12}(t, t_f)Q_f]X(t_f)$, we can eliminate $X(t_f)$ to get $X = [\phi_{11}(t, t_f) + \phi_{12}(t, t_f)Q_f][\phi_{11}(t_0, t_f) + \phi_{12}(t_0, t_f)Q_f]^{-1}X(t_0) = X(t, X_0, t_0)$, now we can find $\lambda(t)$ in terms of $X(t_f)$ as $\lambda(t) = [\phi_{21}(t, t_f) + \phi_{22}(t, t_f)Q_f]X(t_f)$, then we can eliminate $X(t_f)$ to get $\lambda(t) = [\phi_{21}(t, t_f) + \phi_{22}(t, t_f)Q_f][\phi_{11}(t, t_f) + \phi_{12}(t, t_f)Q_f]^{-1}X(t) = \phi_{\lambda x}X(t)$. Now we search a solution for $\phi_{\lambda x}$. By differentiating $\lambda(t)$ we get $\dot{\lambda}(t) = \dot{\phi}_{\lambda x}X(t) + \phi_{\lambda x}\dot{X}(t)$. Comparing the last equation with the Hamiltonian matrix we get $-QX(t) - A^T\lambda(t) = \dot{\phi}_{\lambda x}X(t) + \phi_{\lambda x}\dot{X}(t)$, then we have

$$\begin{aligned} -\dot{\phi_{\lambda x}}(t)X(t) &= QX(t) + A^T\lambda(t) + \phi_{\lambda x}\dot{X}(t) \\ &= QX(t) + A^T\lambda(t) + \phi_{\lambda x}(AX - BR^{-1}B^T\lambda(t)) \\ &= (Q + \phi_{\lambda x}A)X(t) + (A^T - \phi_{\lambda x}BR^{-1}B^T)\lambda(t) \\ &= (Q + \phi_{\lambda x}A)X(t) + (A^T - \phi_{\lambda x}BR^{-1}B^T)\phi_{\lambda x}X(t) \\ &= [Q + \phi_{\lambda x}A + A^T\phi_{\lambda x} - \phi_{\lambda x}BR^{-1}B^T\phi_{\lambda x}]X(t). \end{aligned}$$

Since this is true for arbitrary $X(t)$, $\phi_{\lambda x}$ must satisfy

$$-\dot{\phi}_{\lambda x}(t) = Q + \phi_{\lambda x}A + A^T\phi_{\lambda x} - \phi_{\lambda x}BR^{-1}B^T\phi_{\lambda x}, \tag{1.4}$$

which is the matrix differential Riccati Equation . We can solve for $\phi_{\lambda x}$ by solving Riccati Equation backwards in time from t_f with $\phi_{\lambda x}(t_f) = Q_f$. The optimal control is then given by

$$U^\star = -R^{-1}B^T\lambda(t) = -R^{-1}B^T\phi_{\lambda x}X = -K(t)X(t, X_0, t_0). \tag{1.5}$$

From 1.5 we notice that the optimal control is a linear full-state feedback control, therefore the linear quadratic terminal controller is feedback by default.

2. The Approximating Sequence of Riccati Equations(ASRE)

Assume that we have the following nonlinear system

$$\dot{X} = f(X, U, t) \tag{2.1}$$

$$X(t_0) = X_0, \quad X(t_f) = X_f \in R^n \tag{2.2}$$

with performance index

$$J = \phi(X_f, t_f) + \int_{t_0}^{t_f} L(X, U, t)dt. \tag{2.3}$$

This system can be rewritten in the state dependent quasi-linear system as follows

$$\dot{X}^i = A(X^{i-1})X^i + B(X^{i-1})U^i \tag{2.4}$$
$$X(t_0) = X_0^0, \quad X(t_f) = X_f^n \in R^n \tag{2.5}$$

$$J = {X_f^i}^T Q(X_f^{i-1})X_f^i + \frac{1}{2}\int_{t_0}^{t_f} [{X^i}^T Q(X^{i-1})X^i + {U^i}^T R(X^{i-1})U^i]dt, \tag{2.6}$$

where i represents the iteration step over the time interval $[t_i - 1, t_i]$ Figure 2, the technique is based of the previously introduced Linear Quadratic Regulator with fixed terminal state, which is a full state feedback and therefore the obtained solution will be a closed loop one, i.e. able to respond to the unexpected change in the inputs. The technique works as follows: the initial state is used to compute A_0, and B_0 and we solve for the first LQR iteration and compute X^1 and then used to compute new value of A_1, and B_1 for the second iteration until the final state error reaches a value below a set threshold.

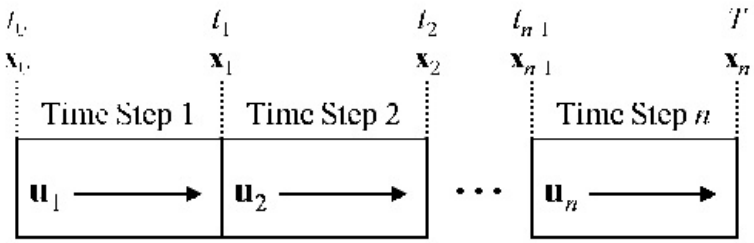


Figure 2. Time Interval Disecretization.

3. Optimal Orbit Transfer

The equations of motion are written in polar coordinates (r, θ), in the inertial Earth-Centered frame. In order to transfer the spacecraft between two circular coplanar orbits two components of the thrust control are used. The tangential component T_θ, and the radial component T_r.

The equations of motion are:

$$\begin{aligned} \ddot{r} - r\dot{\theta}^2 &= T_r - \frac{\mu}{r^2} \\ r\ddot{\theta} + 2\dot{r}\dot{\theta} &= T_\theta \end{aligned} \tag{3.1}$$

where μ is the gravitational constant of the Earth ($3.986005 \times 10^{14} m^3/s^2$) In this system of units the gravitational constant μ is unity, and equations (3.1) are rewritten as:

$$\begin{aligned} \ddot{r} - r\dot{\theta}^2 &= T_r - \frac{1}{r^2} \\ \ddot{\theta} + 2\frac{\dot{r}\dot{\theta}}{r} &= \frac{T_\theta}{r} \end{aligned} \tag{3.2}$$

Equations of motion in state variable form, equations (3.2), are then written in state variable form. The state vector $\mathbf{x}$ is chosen to be:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} r \\ \theta \\ \dot{r} \\ \dot{\theta} \end{bmatrix} \tag{3.3}$$

and the control vector is :

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} T_r \\ T_\theta \end{bmatrix}. \tag{3.4}$$

Then equation (3.2) can be written in the form :

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{B}(\mathbf{x})\mathbf{u}. \tag{3.5}$$

Choosing a suitable factorization equation (3.5) is rewritten in the factored state variable form :

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{x})\mathbf{x} + \mathbf{B}(\mathbf{x})\mathbf{u}, \tag{3.6}$$

where :

$$\mathbf{A}(\mathbf{x}) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ x_4^2 & -\frac{1}{x_1^2 x_2} & 0 & 0 \\ -\frac{2x_4}{x_1^2} & 0 & 0 & 0 \end{bmatrix}, \tag{3.7}$$

$$\mathbf{B}(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & \frac{1}{x_1} \end{bmatrix}. \tag{3.8}$$

4. Factored Controllability

For the factored system (3.6) the controllability is established by verifying that the controllability matrix $\mathbf{M}_{cl} = [\mathbf{B}\ \mathbf{AB}\ \mathbf{A}^2\mathbf{B}\ \mathbf{A}^3\mathbf{B}]$ has a rank equals to $n = 4\ \forall x$ in the domain.

Since $\mathbf{A}$ and $\mathbf{B}$ have nonvanishing rows the controllability matrix $\mathbf{M}_{cl}$ for the System (3.6) is of rank 4.

Nondimensionalization of the problem in order to simplify the calculation we dimensionalize the system by removing the units from the equations of motion via multiplying or dividing some

constants. The two constant we divide by are the radial distance of the initial orbit and the gravitational constant μ in this case the radius of the initial orbit is unity and velocity is divided by the circular velocity of the initial orbit $\sqrt{\frac{\mu}{r_0^2}}$ and the time is multiplied by $\sqrt{\frac{\mu}{r_0^3}}$. In the first two example we would like to make an optimal orbit transfer (i.e. from $(r = 1)$ to $(r = 1.2)$ in time $t_f = 4.469, 5.2231$ (time unit) Figure 3 with an optimal control function of both radial and tangential components Figure 4. The initial angle is $(\theta_0 = \frac{\pi}{2})$ and the final angle is $(\theta_f = \frac{3\pi}{2})$. $\dot{r}_0 = 0$ and $\dot{r}_f = 0$ for the initial and final orbits. $\dot{\theta}_0 = \sqrt{\frac{1}{r_0^3}} = 1$ and $\dot{\theta}_f = \sqrt{\frac{1}{r_f^3}} = 0.54433105395$. In the second $\theta_f = \frac{5\pi}{2}$ with $t_f = 6.866$. In both examples the matrices **Q** and **R** are the identity matrices:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

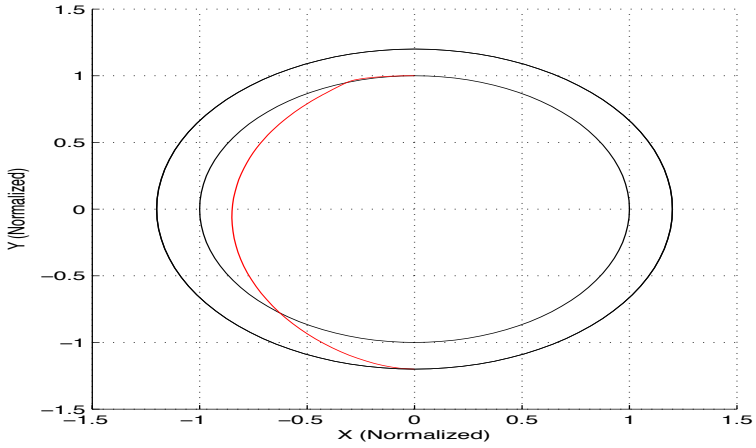


Figure 3. Trajectory of orbit transfer in polar coordinates, from $[r_0 = 1,\ \theta_0 = \pi/2\ , \dot{r}_0 = 0,\ \dot{\theta}_0 = 1]$ to $[r_f = 1.2,\ \theta_f = 3\pi/2,\ \dot{r}_f = 0,\ \dot{\theta}_f = 0.72213]$

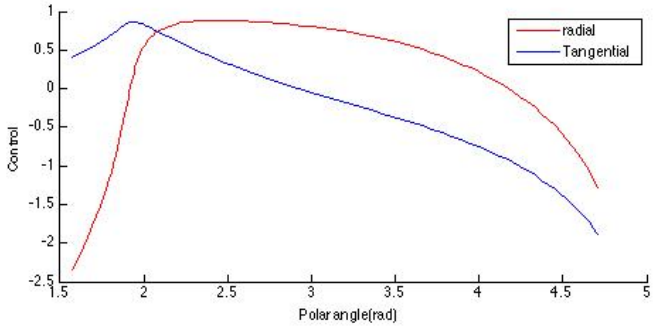


Figure 4. Control function in polar coordinates, from $[r_0 = 1,\ \theta_0 = \pi/2,\ \dot{r}_0 = 0,\ \dot{\theta}_0 = 1]$ to $[r_f = 1.2,\ \theta_f = 3\pi/2, \dot{r}_f = 0,\ \dot{\theta}_f = 0.54433]$.

5. Conclusion

The nonlinear feedback optimal control can be solved by factorizing the original nonlinear dynamics into accessible (weakly controllable) linear dynamics of state dependent factors. The factorized problem has been solved using the the Approximating Sequence Riccati Equations (ASRE) method. The technique is based on Linear Quadratic Regulator (LQR) with fixed terminal state, which guarantees closed loop solution. The method is tested through reconfiguration of a GNSS circular constellation. The result is valid for any circular orbit transfer.

6. Acknowledgments

This project was supported financially by the Science and Technology Development Fund (STDF), Egypt, Grant No 1834.

References

Alizadah, I. and B. F. Villac (2011). Static solutions of the Hamilton-Jacobi-Bellman equation for circular orbit transfer. *Jounal of Guidance, Control and Dynamics* **34**(5), 1584–1588.

Bryson, A. J. and Y. C. Ho (1975). *Applied Optimal Control: Optimization, Estimation and Control*. Taylor and Francis Group. NY.

Chang, D. E. and D. F. Marsden (2002). Lyapunov based transfer between elliptic keplerian orbits. *Discrete and Continuos Dynamical Systems. Series B* **2**, 57–67.

Cimen, T. (2004). Global optimal feedback control for general nonlinear systems with nonquadratic performance criteria. *Systems & Control Letters* **53**(5), 327–346.

Cimen, T. (2006). Recent advances in nonlinear optimal feedback control design. In: *9th WSEAS International Conference on Applied Mathematics, Anonymous ROKETSAN Missiles Industries Inc*.

de Weck, O. L., U. Scialom and A. Siddiqi (2008). Optimal reconfiguration of satellite constellations with the auction algorithm. *Acta Astronautica* **12**, 112–130.

Enright, P. and B. Conway (1992). Discrete approximations to optimal trajectories using direct transcription and nonlinear programming. *Journal of Guidance, Control, and Dynamic* **15**, 994–1002.

Haung, W. (2012). Solving coplanar power-limited orbit transfer probelm by primer vector approximation method. *Internation Journal of Aerospace Engineering*.

Owis, A. (2013). Feedback optimal control of low-thrust orbit transfer in central gravity field. *International Journal of Advanced Computer Science and Applications* **4**(4), 158–162.

Pontryagin, L., V. Boltyanskii, R. Gamkrelidze and E. Mishchenko (1952). *The Mathematical Theory of Optimal Processes*. 2nd ed.. John Wiley & Sons, New York. New York.

Topputo, F., A. Owis and F. Bernelli-Zazzera (2008). Analytical solution of the feedback optimal control for radially accelerated orbits. *Journal of Guidance, Control and Dynamics* **31**(5), 1352–1359.

Topputo, F. and F. Bernelli-Zazzera (2012). A method to solve the nonlinear optimal control problem in astrodynamics. In: *The 1St IAA Conference on Dynamics and Control of Space Systems*. Porto, Portugal.