

利用適應性視窗分析影像紋理之物件分割技術

蔡耀弘

玄奘大學資訊科學所

助理教授

tyh@hcu.edu.tw

呂威廷

玄奘大學資訊科學所

研究生

tim30301717@yahoo.com.tw

摘要

在自然影像中，物件經常存在有大量複雜的紋理，人類經過學習與累積經驗，可以輕易地分辨出不同的物件，目前利用電腦系統來切割物件的方法始終無法達到完全取代人工的目的。在著名的灰階共現矩陣(Gary level co-occurrence matrix, GLCM)方法中，作者藉由固定大小的視窗掃描整張影像，利用統計的方法，計算視窗內的紋理特徵，藉此分割出具有不同紋理的物件。但是這個方法存在著固定視窗大小無法準確分割影像中不同大小紋理的問題。本篇論文針對此問題提出改良，利用適應性視窗來分析影像紋理，設計有效的演算法，找出影像中每個區域最適合的視窗大小來做統計特徵分析。在分析過後得到統計特徵矩陣，接著我們利用 K-means 的分群法來對此矩陣做分類並得到物件分割的結果。

關鍵詞：影像紋理分析、適應性視窗、灰階共現矩陣、物件切割

1. 前言

影像處理、電腦視覺的應用，在近年越來越常見，例如：數位相機人臉追蹤功能、監視系統、人臉辨識、車牌辨識、人臉追蹤、影像檢索……等。為了能夠支援上述的各種應用，通常會先將一張影像或一段影片中的不同物件分割出來，其結果再視不同應用做不同方法的處理[3]。目前常見的物件切割方法，有基於色彩(color-base)、基於形狀(shape-base)、基於模型(model-base)、特徵訓練、紋理分析(texture analysis)等多種方法[4,7,12]。由於自然影像中的紋理非常複雜，大多數物件的紋理也沒有一定的規則與定律，目前大多的分析方法只能透過統計的方式，來預測出物件的紋理特性。基於紋理的不規則性，相較於上述其他幾種物件切割方法，在電腦的分析處理上是最困難的，但也是最有效的方法之一。

在自然影像中，不同物件會存在不同紋理，從自然的樹皮、天空、草原到人為的馬路、大樓、衣服等。而在許多領域中，例如：醫療診斷影像、瑕疵品檢測、衛星影像、地質研究等領域，在以往都需要依靠專業人士大量的人工分析，將不同物件分割出來，以利後續處理作業。因此，在紋理分析的課題中，希望能夠藉由電腦系統來判斷紋理，將不同紋理的不同物件分割出來，以此減輕人工負擔。

本論文針對紋理分析(texture analysis)中著名的二階灰階分析法—灰階共現矩陣(Gary level co-occurrence matrix, GLCM)來提出改良[11]。GLCM 是利用一固定大小的視窗對影像做掃描，並逐一計算每個視窗內各角度間兩相鄰像素點的各種灰階值組合機率，接著利用各種統計方法來計算其關聯性，藉由不同紋理特徵來分辨不同物件，進而達到分割的效果。

GLCM 存在著固定的視窗大小無法對不同影像紋理、物件大小做準確的切割，必須依照情況調整視窗的大小。此外，假如對自然影像來做物件分割，經常會遇到一張影像中同時存在不同大小的物件、紋理，固定的視窗大小就無法準確合適的做切割。視窗過大會將小物件視為紋理的一部分而忽略，而太小的視窗則會將紋理中的變化誤判為不同的物件。Marceau[6]指出，在使用 GLCM 來分析紋理時，影像解析度、灰階值量化數、視窗大小、角度、特徵統計量等參數都會影響紋理分析的準確度，而視窗大小對影像的影響程度則占了高達 90%的比例。Rishi Jobanputra 與 David A. Clausi[9][10]提出一個在 GLCM 計算每個視窗中的灰階值組合機率時，加入權重的方法 Weighted GLCM(WGLCM)，此方法在處理單純紋理影像時可以得到比 GLCM 較佳的物件切割結果，但是無法有效處理自然影像中不同尺寸的影像紋理分割。

我們提出了一個利用適應性視窗分析影像紋理之物件分割的技術，來改善既有的 GLCM 中因為固定視窗大小所造成的問題。我們提出的方法會隨分析結果自動調整視窗大

小，讓不同區域的統計特徵能夠利用不同視窗大小，得到較適合的分析結果，藉此增強物件切割準確度。

2. GLCM

紋理(texture)是在影像中物件的灰階值變化機率與幾何分佈，或是由某種規律的圖案所組成排列而成。

常見的紋理分析方法有兩類，一種是統計方法(statistical methods)，一種是結構方法(structural methods)[8]。統計方法是利用紋理在影像空間上的灰階分配特性，找出其中的統計關聯性，藉此區分出不同統計特徵的物件紋理。另一種結構方法，則是找出紋理中重複排列組成的圖案，利用其規律性的結構來分割出不同物件。由於自然影像中，所涵蓋的物件紋理非常複雜，極少會有由規律的圖案所組成的物件，因此較不適合使用結構方法來分割物件。我們所改良的 GLCM 是紋理分析中統計方法的一種，也是較適合分析自然影像的一種方法。

灰階共現矩陣(Gray Level Co-occurrence matrix, GLCM)，是由 Haralic[11]等人於 1973 年所提出的。它的作法是先對一張影像的灰階值作量化，然後設定一正方形且固定的視窗大小，視窗由左至右，由上到下，每次移動一個像素的距離，將一張影像逐一掃描過。在掃描的過程中，記錄每個視窗中各種灰階值的組合，並計算條件聯合機率，再依需求使用不同統計方法來計算統計特徵。詳細方法說明如下：

2.1 灰階值量化

我們假設灰階值量化的數量為 G ，我們先對一張影像作量化。圖 2.1 是對一個 5×5 大小的影像將灰階值從 16 色量化為 4 色之後的結果。

G. Soh 與 Tsatsoulis[5]指出，將影像中的灰階值量化到 64 就足夠。Clausi[2]認為 G 超過 64 是過大的。過大的 G 將會導致過大的矩陣，以及大量的 0 組合次數。一般會將影像量化在數量 16 到 64 間，本論文將會以 $G=32$ 作為實驗時的參數設定。

定義的方法如圖 2.2 所示，設當前的像素點為 i ，相鄰的像素點為 j 。圖例中說明距離， $\delta=1$ 與角度， $\theta=90^\circ$ ，及距離， $\delta=2$ 與角度， $\theta=0^\circ$ 時的配對方式。

2	6	12	11	5
8	15	7	5	3
13	9	7	12	4
15	7	0	9	2
6	11	12	7	3

(a)

0	1	3	2	1
2	3	1	1	0
3	2	1	3	1
3	1	0	2	0
1	2	3	1	0

(b)

圖 2.1、視窗大小 5×5 ，(a) $G=16$ ，(b) $G=4$ 。

2.2 定義像素組的距離與角度

GLCM 會計算出視窗內成對像素點的各種灰階值組合機率。在此之前，我們必須先定義出像素組之間的距離(δ)與角度(θ)關係。一般常用的角度有： 0° 、 45° 、 90° 、 135° 四個方向，主要原因是方便計算，且這四個角度也包含了水平、垂直及兩個對角線的紋理方向。

$$(i, j, \delta, \theta) = (0, 1, 1, 0^\circ)$$

0	1	3	2	1
2	3	1	1	0
3	2	1	3	1
3	1	0	2	0
1	2	3	1	0

(i, j, δ , θ) = (0, 1, 2, 90°)

圖 2.2、像素組的配對方式說明。

2.3 灰階共現矩陣

經過定義距離與角度後，我們根據此定義，對之前量化後的矩陣依序將視窗中兩兩像素點間灰階值的出現頻率記錄到矩陣中。

執行方式為：先根據量化數 G 產生一個 $G \times G$ 的空矩陣，列座標代表當前基準像素點的灰階值 i ，行座標則表示相鄰點的灰階值 j 。矩陣中的值所相對應的座標，則代表視窗中當基準點的灰階值 i 與相鄰點的灰階值 j 所產生的次數。

如圖 2.3，我們以距離(δ)為 1，角度(θ)為 0° ，量化數(G)等於 4 為例，將圖 2.1 中的像素組灰階值出現次數紀錄到矩陣中。

接著，我們透過公式 2.1，將圖 2.3 的矩陣轉換為條件聯合機率(conditional joint probabilities)，如圖 2.4。

P_{ij} 代表視窗中基準點灰階值 i ，與相鄰點

灰階值 j 的所共同產生的次數，即為圖 2.3 中座標 (i,j) 中的值。 G 為量化數。公式 2.1 的下半部為一個視窗中 i 與 j 所有可能產生的組合總次數。 C_{ij} 則為條件聯合機率(conditional joint probabilities)。

		j			
		0	1	2	3
i	0	0	1	1	0
	1	3	1	1	2
	2	1	2	0	2
	3	0	4	2	0

$i=3, j=1$ 的產生次數總共為 4

圖 2.3、 $G=4$ ， $(\delta, \theta)=(1, 0^\circ)$ 的條件下產生的灰階共現矩陣。

$$C_{ij} = \frac{P_{ij}}{\sum_{i,j=0}^{G-1} P_{ij}} \quad \text{公式 2.1}$$

		j			
		0	1	2	3
i	0	0/20	1/20	1/20	0/20
	1	3/20	1/20	1/20	2/20
	2	1/20	2/20	0/20	2/20
	3	0/20	4/20	2/20	0/20

圖 2.4、將圖 2.3 中的矩陣轉換為條件聯合機率後的矩陣。

2.4 統計特徵

在紀錄出灰階共現矩陣後，會依據不同的需求，套用不同的統計方法來求出統計特徵。表 2.1[9]列出常見的 8 種統計方法。

Baraldi 與 Parmiggiani[1]認為 energy 以及 contrast 兩個統計方法是比較適合來做紋理統計分析的，我們實驗的過程中，也發現 energy 的統計方法在最後利用 K-means 做分割時，所得到的結果較佳。

3. 我們的方法

本篇論文針對灰階共現矩陣(Gray Level Co-occurrence Matrix, GLCM)方法中固定視窗大小導致的問題來提出改良。圖 3.1 為我們提出改良方法的流程圖，在後面的小節將會對我們所提出的方法做詳細的說明。

表 2.1、常見的統計方法與公式。

統計方法	公式
Maximum probability(MAX)	$\max(C_{ij})$ for all (i,j)
Uniformity(UNI)	$\sum C_{ij}^2$
Entropy(ENT)	$\sum C_{ij} \log C_{ij}$
Dissimilarity(DIS)	$\sum C_{ij} \cdot i - j $
Contrast(CON)	$\sum C_{ij} (i - j)^2$
Inverse difference(INV)	$\sum \frac{C_{ij}}{1 + i - j }$
Inverse difference moment(IDM)	$\sum \frac{C_{ij}}{1 + (i - j)^2}$
Correlation(COR)	$\frac{\sum (i - \mu_x)(j - \mu_y)C_{ij}}{\sigma_x \sigma_y}$

3.1 灰階共現矩陣(Gray Level Co-occurrence matrix, GLCM)

一開始，我們先利用前一章節所提到的灰階共現矩陣(GLCM)對整張影像做一次分析，然後得到一張統計特徵的矩陣。必須注意的是，一開始的視窗必須使用較大的尺寸，視窗大小在每次執行過後減少 2 個像素寬。在本篇論文的實驗中，我們設定初始的視窗大小為一張影像最小邊長的一半，像素組的距離(δ)為 1，角度(θ)為 0° ，量化數(G)為 32。

3.2 找出轉折點

在經過第一次 GLCM 的統計後，我們會得到統計特徵的矩陣。所謂轉折點，在這篇論文的定義是，當我們將矩陣中某一列數值畫成直方圖時，在這條線上的所有方向轉換的點，我們稱為轉折點。而每兩個相鄰的轉折點之間所有的點，則構成一個區段。接著，我們先逐列的掃描整張矩陣，找出矩陣中所有的轉折點。

如圖 3.2 所示，黑點部份即為轉折點。需要注意的是，起始點與結束點要預設為轉折點，如此在下一步計算斜率的時候，才不會將頭尾兩區段給忽略掉。

找出水平方向的所有轉折點後，我們在以此方式找出垂直方向的所有轉折點，但是水平方向與垂直方向的轉折點必須區別，分開紀錄。

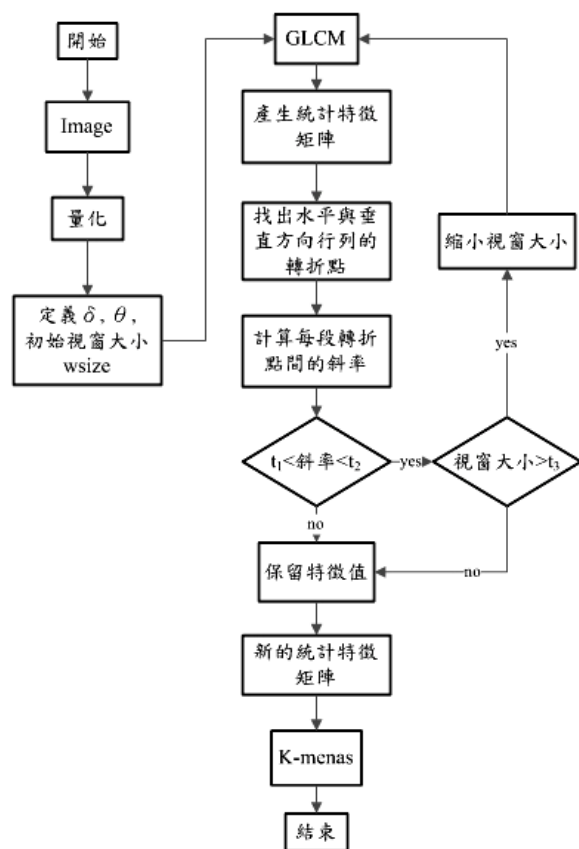


圖 3.1、利用適應性視窗分析影像紋理之物件分割技術之流程圖。

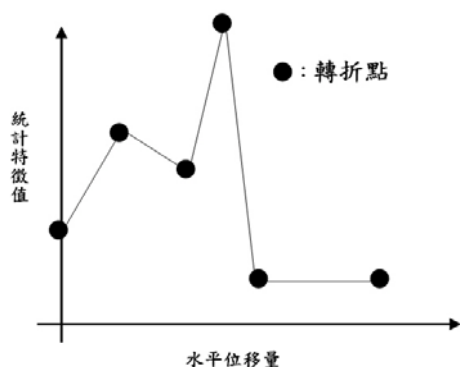


圖 3.2、轉折點示意圖。

3.3 計算每個區段的斜率

以水平方向所紀錄的轉折點來看，我們定義每個左右相鄰的轉折點間所有的點所構成的線段為一區段，反之，垂直方向亦然。

一般來說，一張複雜影像所產生的統計特

徵矩陣，每一列或行所形成的直方圖都會有相當多的震盪與雜訊。因此，假如所有大小震盪的轉折點我們都列入考慮，不僅會造成計算上沉重的負擔，也會產生許多錯誤。另一方面，某些紋理的特性是有規律的震盪，並不需強化它來加強切割。所以，我們應該忽視某種程度的震盪，僅針對較劇烈的起伏做計算。圖 3.3(a)(b)說明消除震盪的過程與結果。

因此，在我們的實驗中處理震盪的方法，是將上下震幅不到縱座標百分之十的區段給忽略掉，根據我們的實驗，大部分紋理中的小震盪都不會超過這個範圍。

由於在方法的最後，我們會以 K-means 的分群法來對矩陣做切割，假如在區段間的變化愈劇烈，也就是紋理分界處落差越大越急，則在 K-means 的切割上就愈精確。

所以，我們希望保留斜率極低（接近水平）以及斜率極高（接近垂直）的區段。而斜率介於極高與極低之兩門檻值中間的區段（較平緩提升或降低），我們將會取出，重新以較小的視窗計算此區段間每一點的統計特徵，此區段則稱為我們的目標區段。

圖 3.3 表示找出所有轉折點，消除震盪與雜訊以後，取出目標區段，並且對該區段的值以較小的視窗重新計算後的結果。

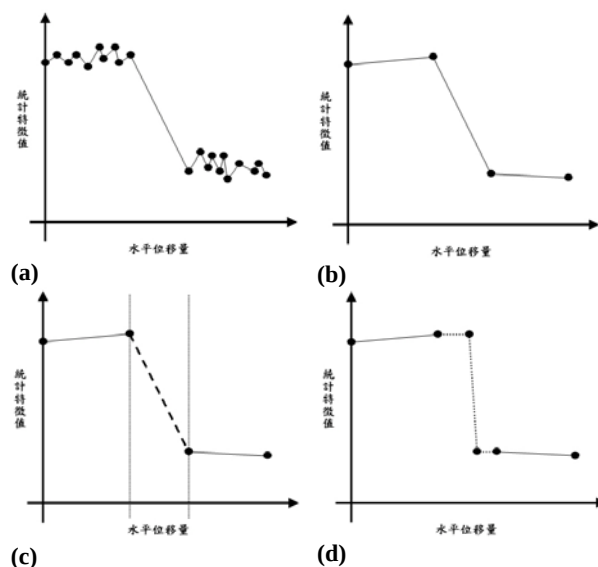


圖 3.3、(a)找出所有轉折點，(b)消除線段中的小震盪，(c)取出斜率介於 t_1 與 t_2 間的目標區段，(d)將取出的區段以較小的視窗計算特徵後的結果。

3.4 調整視窗大小

我們先用較大的視窗，可以先得到較為概括的分析，通常會產生一個較為平滑無震盪的

直方圖，因為紋理中的細小變化會被大的視窗給忽略，也降低之後分割時產生雜訊的情形。假設使用較小的視窗，會造成較為劇烈的變化，紋理中的小變化會被小視窗所突顯。因此我們取出介於極高與極低門檻值之間的目標區段，將之以較小的視窗重新計算後，會得到較為劇烈的變化，藉此將紋理交界處的變化突顯出來。

3.5 重複運算

我們的方法會在執行一次 GLCM 後，就檢查整張矩陣在水平方向或垂直方向上是否需要重新計算的區段，假如有，就對該區段中所有的點使用縮小的視窗重新計算其統計特徵，並加以更新至原來的統計特徵矩陣中，在對此新的矩陣找出斜率未達門檻值的區段，再重複 GLCM 的階段。也就是不斷重複章節 3.1~3.4 的步驟，直到整個統計特徵的矩陣沒有需要重新計算的區段，或是視窗大小已小於 7×7 時，就停止計算得到收斂結果，最後將產生的統計特徵矩陣則繼續下一階段的 K-means 分群法來做切割。

3.6 切割 (Segmentation)

在 3.4 節中最後得到的統計特徵矩陣，我們使用 K-means 的分群法來對此矩陣做分割。

所謂 K-means 的方法，是先決定要分群的數量(k)，然後對要分群的資料隨機灑下 k 個種子點，接著以這 k 個點的值對整筆資料作叢集，並在資料中產生 k 個分群。在初步分群過後，分別計算這 k 個群的平均值，再以這 k 個平均值作分群。不斷重複上述分群、求平均值、分群的步驟，反覆直到結果不變即為收斂。最後，不同群的區域間的交界處即為我們的分割線。

因為 K-means 在一開始灑點是利用隨機的方式，因此最後的結果可能會有些微不同，但是不會有太大的差異。

雖然直接使用 K-means 分群法也能夠直接對影像灰階值來做切割，但是我們的方法針對紋理分析後的統計特徵矩陣來做切割，所得到的結果會更接近人類對物件的認知與分辨。

圖 3.4 表示從原圖到 GLCM 所產生的每列直方圖，然後利用我們的方法處理過後使用 K-means 做分群的結果。

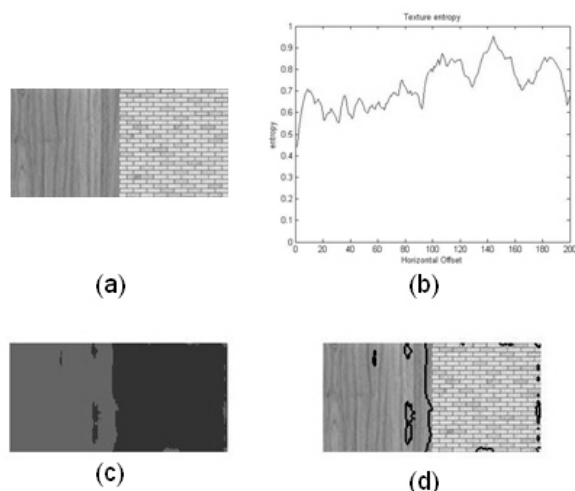


圖 3.4、(a)原圖，(b)統計特徵矩陣其中一列所產生的直方圖，(c)利用 K-means 對統計矩陣做分群，(d)分割結果

4. 實驗結果

4.1 參數設定

在我們的實驗中，設定量化數 $G=32$ ，初始的視窗尺寸為影像最小邊長的一半，視窗大小的邊長每次減少 2 個像素，最小縮小到 7×7 大小的視窗便停止。我們使用的 GLCM 參數為 $\delta=1$ 與 $\theta=0^\circ$ 。再決定目標區段的門檻值 t_1 與 t_2 ，我們則設定 t_1 為 0.5， t_2 為 15。需注意的是，我們在執行 GLCM 時所產生的統計特徵矩陣，已經經過正規化的換算過程，因此矩陣中的值會介於 0 到 1 之間，垂直區段的斜率就不會過高，在我們實驗中，介於斜率 0.5 到 15 之間，已經能將目標區段給擷取出來。在統計方法上，我們使用較適合做影像紋理分析的 energy 方法來計算特徵。表 4.1 列出我們實驗中預設不變的參數與參數值。

表 4.1、實驗預設參數值

參數	設定值
G	32
δ, θ	1, 0°
t_1, t_2	0.5, 15
Statistic method	energy

4.2 比較 GLCM 與我們的方法

圖 4.1 表示 GLCM 與我們的方法在紋理影像上的分割結果，影像大小為 100×50 ，圖中的黑線或白線為不同紋理的分割線，分割線與

實際紋理分界處若越吻合，即分割結果越為準。實驗結果顯示我們方法的分割線較 GLCM 的方法更為準確。圖 4.2 表示了利用我們的方法與 GLCM 相對照處理自然影像的結果，圖 4.2(a)是我們輸入的影像，圖 4.2(b)是使用 GLCM 處理的結果，圖 4.2(c)顯示利用我們提出的方法來分割影像物件的結果。在圖 4.2 中，我們依上、中、下順序對樹木、雲、森林與建築物來做紋理分割，圖 4.2(b)與圖 4.2(c)相比較下，能夠明顯看出後者的分割線較貼近實際物件的分界線，在中間的圖裡，右上方散落的小雲也都能夠準確切割出來，而下圖中建築物、森林與天空的分界處也較 GLCM 的方法更精準。



圖 4.1、(a)GLCM，(b)我們的方法。

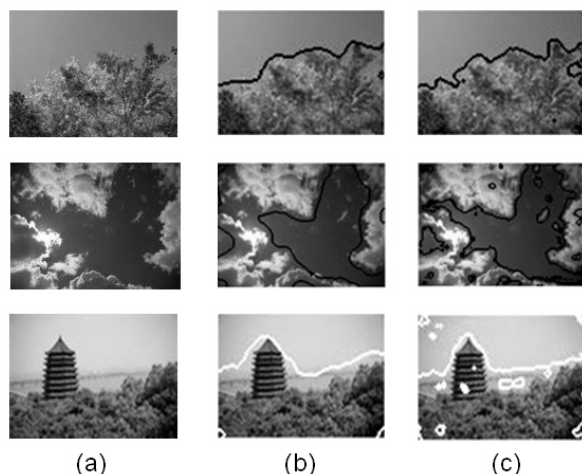


圖 4.2、(a)原圖，(b) GLCM，(c)我們的方法。

5. 結論

我們的方法，主要針對 GLCM 中視窗大小所產生的問題提出改良。在一般紋理影像中，或許 GLCM 可以經由測試找出較佳的視窗大小來作分析，但是處理自然影像中涵蓋大小不同的物件與紋理時，就無法以單一視窗大小去切割不同物件。而我們的方法可以藉由反覆計算，來找出適用於不同區域的不同視窗大小，使較小的物件不會因為過大的視窗被忽略，較大的物件也不會因為太小的視窗導致其中的紋理被錯誤辨識小物件。

我們的實驗結果在某些區域有錯誤分割的情形產生，主要是門檻值調整的問題，但是這些門檻值的影響並不會大於視窗大小的影響。因此，我們的方法除了能夠自動找出較佳的視窗大小外，還允許不同區域使用不同視窗大小進行分析，解決了 GLCM 單一視窗大小的根本問題。

6. 參考文獻

- [1] A. Baraldi, F. Parmiggiani, An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters, *IEEE Trans. Geosci. Remote Sens.* 33 (2) (March 1995) 293-303.
- [2] D.A. Clausi, An analysis of co-occurrence texture statistics as a function of grey level quantization, *Canad. J. Remote Sens.* 28 (1) (2002) 45-62.
- [3] Gonzalez Woods, "What is image processing?" Digital Image Processing 2/e, Chapter 1.1, Prentice Hall, 2004.
- [4] Haim Permuter, Joseph Francos, Ian Jermyn, "A study of Gaussian mixture models of color and texture feature for image classification and segmentation." *Pattern Recognition* 39, 695-706, 2006.
- [5] L.K. Soh, C. Tsatsoulis, Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices, *IEEE Trans. Geosci. Remote Sens.* 37 (2) (1999) 780-794.
- [6] Marceau, D., Howarth, P. J., and Dubois, Jean-Marie M., "Automated Texture Extraction from High Spatial Resolution Satellite Imagery for Land-Cover Classification," *International Geoscience and Remote Sensing Symposium*, pp. 2765-2768, 1989.
- [7] Mats Erikson, "Two preprocessing techniques based on grey level and geometric thickness to improve segmentation results," *Pattern Recognition Letters*, 27, 160-166, 2006.
- [8] M. Tuceryan, A.K. Jain, Handbook of Pattern Recognition and Computer Vision, Chapter 2: Texture Analysis, World Scientific, Singapore, 1993.
- [9] Rishi Jobanputra, David A. Clausi, "Texture Analysis Using Gaussian Weighted Grey Level Co-occurrence Probabilities," *Computer society*, 2004.
- [10] Rishi Jobanputra, David A. Clausi, "Preserving boundaries for image texture segmentation using grey level co-occurring probabilities," *Pattern Recognition* 39 234-245, 2006.

- [11] R.M. Haralick, K. Shanmugam, I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cyber.* 3 (6), 610-621, 1973.
- [12] Yining Deng, B.S. Manjunath, "Unsupervised Segmentation of Color-Texture Regions in Images and Video," *IEEE Transactions on pattern analysis and machine intelligence*, VOL. 23, NO.8, 2001.