تصنيف الصورة الرقمية باستخدام خوارزمية العنقدة

*أحمد حمود فليح ** مؤيد عبد هاني مخور * رائد عباس صالح
* جامعة ذي قار/ كلية العلوم/ قسم الفيزياء
**جامعة ذي قار/ كلية التربية/ قسم الحاسبات

الخلاصة

اعتمدت الدراسة في هذا البحث على طريقة التصنيف غير الموجة للصورة الرقمية باستخدام خوارزمية العنقدة، وتم اقتراح تخفيض الإختيار العشوائي للمتجهات باستخدام عدد العناقيد غير الصفرية في مخطط العلاقة بين التدرجات الرمادية للمتجهات وتكرار ها، كذلك تم استخدم كاشف الحافة لتقليل الأخطاء التي تنشأ من اختيار مناطق الحافات كأصناف طيفية.

١.مقدمة Introduction

هناك محاولات عديدة لتصنيف الصورة عن طريق استخدام الشدة المسجلة لعناصرها حيث يتم مطابقه مجموعة من العناصر المتجانسة والتي تمثل اصنافها المميزة، باعتماد معلوماتها الطيفية الممثلة رقمياً لحزمة طيفيه واحده أو أكثر. يطلق على هذا النوع من التصنيف بتميّز النمط الطيفي Spectral pattern recognition ، إذ يتم تمييز كل عناصر الصوره الرقمية الى مجموعة من الأصناف المحددة[١].

٢. التنصيف الموجه وغير الموجه Supervised and Unsupervised classification

يتضمن التصنيف مدى واسع من طرق إتخاذ القرار النظري في تشخيص الصورة (أو جزء منها). أن جميع خوارزميات التصنيف تستند على افتراض إن الصورة لمناطقها المختلفة سمات مختلفة والذي من الممكن إن نميز كل صنف من سماته. وهذه السمات تختلف باختلاف الأهداف الموجودة في الصورة واختلاف الصور، فمنها السمات المعتمدة على الشدة وتغاير الشدة الرمادية وسمات لونية للصورة. إن التصنيف يقسم إلى نوعين من التصنيف؛ الأول يعرف بالتصنيف الموجة والذي يعتمد على استقطاع من الصورة المختلفة وثم تحدد سماتها حيث تجرى عملية التصنيف بالاعتماد على طريقة المسافة الدنيا أو الاحتمال الأعظم، والنوع الآخر يعرف بالتصنيف غير الموجة وهذا يعتمد على استخلاص سمات مناطق الصورة بشكل آلي باستخدام طريقة حساب العناقيد ليجرى بعد ذلك عملية التصنيف[٢].

٣. خوارزمية العنقدة The Cluster Algorithm

هنالك العديد من طرق التصنيف المستخدمة في التحسس النائي، ومن أكثر الطرق شيوعاً هي خوارزمية العنقدة، والتي تستند على مجموعة من إجراءات تكرارية تتلخص بتحديد عدد من المتجهات العشوائية، ومن ثم يتم تحديد المجاور الأقرب عن طريق حساب المسافة بين المتجهات وعناصر الصورة وبالتالي فان كل متجه قد صنف إلى المجاور الأقرب لهNearest neighbor ، بعد ذلك يتم تحديد عناقيد جديدة عن طريق حساب معدل كل العناصر الأقرب لكل صنف. تكرر هذه الخطوات حتى تقل المسافة عن عتبه محدده مسبقاً أو يحدد التكرار بعدد معين[٣] .

٤.كشف الحافات Edge Detection

الحافة هي وصف لحدود وحواف الأجسام والأهداف الموجودة في الصورة[٣] وهي تمثل فرقاً واضحاً في الشدة أو النسيج وعملية كشف الحواف تعتمد أساسا على عدم الاستمرارية في دالة التجانس لصفات مناطق الصورة [٤,٥].
وكشف الحافات عامل مهم في تحليل واستنتاج المعلومات المفيدة في الصورة، وقد استخدمت العديد من مؤثرات كشف الحافات لتطبيقها على الصور المختلفة لغرض تحليلها وهي عبارة عن مرشحات تستخدم نوافذ لأبعاد مختلفة مثل (٣ x3 ،5x5 ..)ومن ثم نحسب قيمة مجموع العمليات في النافذة والتي قد تكون ناتجة من الالتفاف الرياضي أو ناتجة من عمليات رياضية أخرى باستنتاج قيمة موقعيه يتم مقارنتها مع عتبة محددة لتحديد الحافات في الصورة.

٥. الأصناف المعلوماتية والطيفية The spectral and information classes

إن عملية التصنيف تتطلب من المحلل التمييز بين أصناف المعلومات والأصناف الطيفية، فأصناف المعلومات هي تلك الأصناف ذات الأهمية للمحلل والذي يحاول تعرَّيفها في الصورة، فعلى سبيل المثال غابات مختلفة بالنوع أو أشجار بأنواع خاصة أو أنواع مختلفة من الصخور. أما الأصناف الطيفية فهي مجاميع من العناصر الرقمية للصورة والتي تكون متجانسة في السطوع[١].

أن دور المحلل هو تقرير أهمية اختلاف الطيف للأصناف وانسجامها مع أصناف المعلومات المهمة، ففي التصنيف الموجة فأن المحلل يحاول أن يُحدد من الصورة بنماذج متجانسة كعينات للمناطق المختلفة والتي تغطي الفئات أو الأنواع المهمة. هذه العينات تمثل مناطق تدريب train area وان اختيارها يعتمد على المحلل ومعرفته بحقيقة هذه الأسطح أوالمناطق، وبذلك فأن المحلل يحدد المشهد إلى مجموعة محدده من الأصناف. ويتم تمييز الأصناف حاسوبياً بعد تحديد المعلومات الرقمية في الحزم الطيفية للعناصر التي تؤلف المناطق وحساب التقارب بين النماذج الطيفية المحددة وعناصر المؤلفة للصورة، وبهذا فأن هذا النوع من التصنيف يجب أولا تحديد أصناف المعلومات ومن ثم حساب الأصناف الطيفية. ويكون الحال معاكساً في حالة التصنيف غير الموجه، فالأصناف الطيفية تصنف أولاً ومن ثم المطابقة بواسطة المحلل إلى أصنافها المعلوماتية[٦].

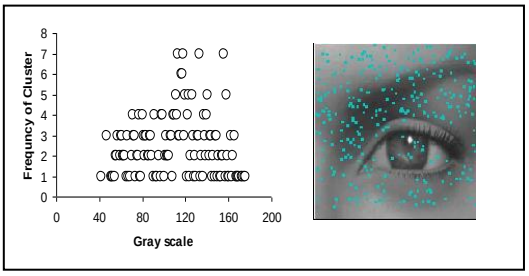
٦. النتائج والمناقشة Results and discussion

تمحور البحث على التصنيف غير الموجة للصورة الرقمية، وقد افترضت طريقة جديدة لتخفيض عدد العناقيد باستخدام المخطط التكراري (histogram) وتتلخص الطريقة بما يأتي:

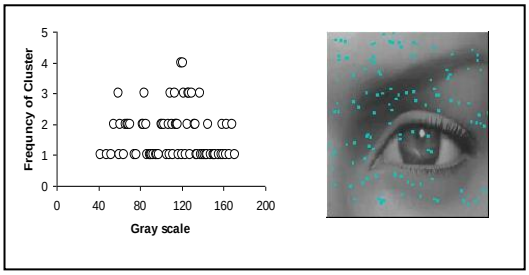
١- اختيار متجهات عشوائي لعدد الكتل البيانية بحجم (٣*٣) عنصر.
٢-حساب المعدل لهذه الكتل البيانية ومن ثم حساب المخطط التكراري لهذه المتجهات.
٣-حساب المسافة الأقرب باستخدام خوارزمية المسافة الدنيا بين متجهات العشوائية وعناصر الصورة الرقمية لينتج عنه أعادة بناء الصورة بتدرجات طيفية جديدة تنشأ من حساب معدل عناصر الصورة الرقمية القريبة من المتجهات العشوائية،
٤- تكرر الخطوات المذكورة أعلاه ولكن باختيار عشوائي لعدد العناقيد غير صفرية في مخطط العلاقة بين التدرجات الرمادية للمتجهات وتكرار ها.

أن اختيار الكتل البيانية يجب أن يقتصر تمثيله في المناطق المتجانسة فقط وكما هو الحال في التصنيف الموجه لذا أقترح استخدام كاشف الحافة لتحدد المواضع أو المواقع التي يحصل فيها تغير في اللون أو الشدة وقد استخدمنا طريقة المربعات الصغرى Liner Square fitting في أيجاد أفضل خط مستقيم وهذه الطريقة تعتبر من الطرق الكفؤة في كشف الحافات، إذ تمتاز المناطق المتجانسة بخط أفقي يكون التغاير عن المعدل أقل فيما كانت المنطقة تمثل حافة، يقارن الجمع التراكمي للتغاير عن المعدل في هذه الطريقة بعتبة معينة ولنافذة محددة (تم استخدام نافذة (٣*٣) وعتبة(Threshold=120)) وكانت النتائج العملية كما يلي:

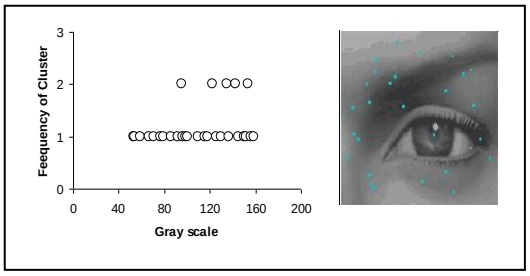
شكل (١) يوضح الصورة الأصلية وكاشف الحافة لها.



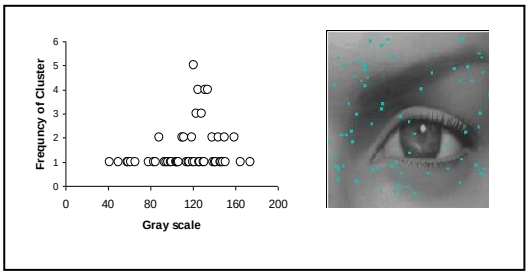
شكل (٢) يوضح الإختيار العشوائي والمخطط التكراري لعناقيد كتل البيانات العشوائية وبحجم (٣*٣ pixel).



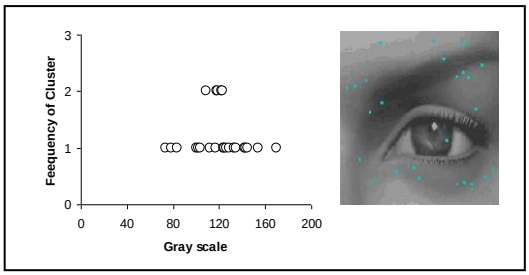
شكل (٣) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



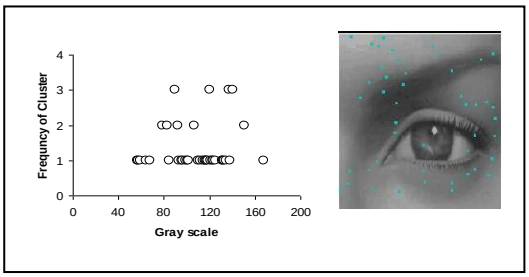
شكل (٧) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



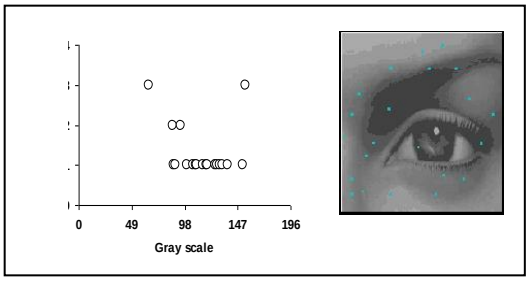
شكل (٤) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



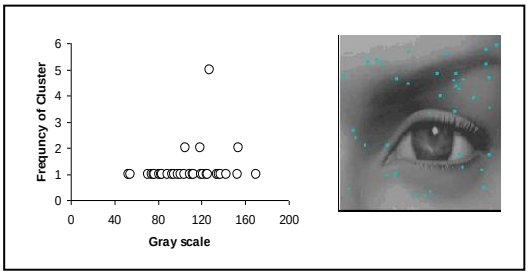
شكل (٨) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



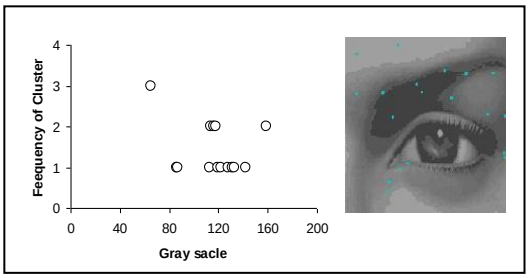
شكل (٥) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



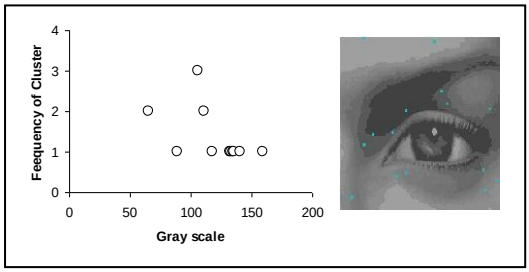
شكل (٩) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



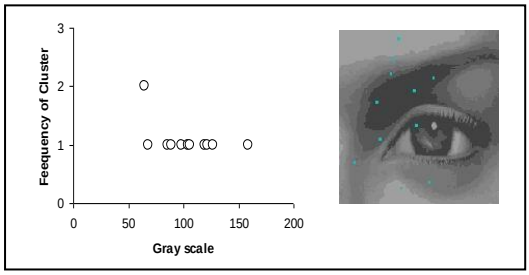
شكل (٦) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



شكل (١٠) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.

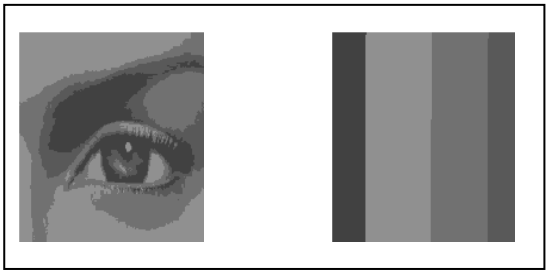


شكل (١١) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.



شكل (١٢) عملية التخفيض في عدد كتل البيانات العشوائية والمخطط التكراري لعناقيدها.

وإن الصورة النهائية والمؤلفة من أربع أصناف لعملية تصنيف الصورة الرقمية باستخدام خوارزمية العنقدة ونسبة التدرجات الرمادية للصورة والتي تعكس المخطط التكراري موضحة بالشكل الآتي:



الشكل(١٣) يوضح تصنيف الصورة لأربعة أصناف ونسب التدرجات الرمادية.

Reference

[1] Fundamentals of Remote Sensing, http://www.nrcan.gc.ca/notice_e.html, 2004

[2] JENSEN, "Unsupervised Classification algorithms.htm", 1996

[3] Charkreyavanich, U., "Edge Detection Operators. Technical report", No.418, Department Geodetic Science and Sarveying, ohio state University, pp. 17. 27,1991.

[4] Digital Imge Processing Documentation-http://documents

wolfram.com/applications/digitalimage/users Guide/7.3.htme.

[5]D.paulus "EDGE IMAGES" CHAPTER 14.pdf, 2001.

[6] https://www.e-education.psu.edu/courses/geog5121/L08_IIIC2.html, 2003