Split and Merge Regions of Satellite Images using the Non-Hierarchical Algorithm of Cluster Analysis

شطر ودمج المناطق لصور الاقمار الصناعية بأستخدام خوارزمية التحليل العنقود غير الهرمية

أ.م.د. اسماء غالب الراوي / كلية الادارة والاقتصاد / جامعة بغداد
drasmaa.ghalib@coadec.uobaghdad.edu.iq
الباحث / محمد عبد الودود محمد a.mohamed@coadec.uobaghdad.edu.iq

OPEN ACCESS
P - ISSN 2518 - 5764
E - ISSN 2227 - 703X

Received:4/9/2018
Accepted: 22/10/2018

المستخلص:

يعد التقطيع الصوري من الاهداف الرئيسة والضرورية في المعالجات الصورية للصور الرقمية، فهو يسعى الى تجزئة الصور المدروسة الى مناطق متعددة اكثر نفعاً تلخص فيها المناطق ذات الافادة لصور الاقمار الصناعية، وهي صور متعددة الاطياف ومجهزة من الاقمار الصناعية باستخدام مبدأ الاستشعار عن بعد والذي اصبح من المفاهيم المهمة التي تُعتمد تطبيقاته في اغلب ضروريات الحياة اليومية، وخاصة بعد التطورات المتسارعة التي شهدتها ميادين الحياة المختلفة والتي كثيراً منها طرقت بابها خوارزميات وتقنيات البرمجيات، فهذه الصور تعد ضرورية جداً لتمكيننا من دراسة طيف واسع من الاهداف في العديد من الجوانب العلمية، في هذا البحث استخدمت خوارزمية التحليل العنقودي غير الهرمية كطريقة للتقطيع الصوري (شطر ودمج المناطق) بهدف عرض اهمية استخدام الاساليب الاحصائية في مهام المعالجة الصورية مثل التقطيع الصوري، حيث اعتمد على تقنية (K-Means) لتنفيذ هذه المهمة، وقد طبقت خوارزمية هذه التقنية على صورة اقمار صناعية متعددة الاطياف لمشهد غربي العراق، حيث اظهرت النتائج مدى مرونة هذه الخوارزمية في التعامل مع التفاوت في اضاءة العناصر الصورية للصورة الملونة وكفاءة تكوينها لمناطق العناقيد المتكونة من مجاميع من العناصر الصورية المتجانسة في درجة شدة اضاءتها، واخيراً قدرة هذه الخوارزمية على اعطاء صور تتميز بجودتها والتي قيست على وفق مقياس ارتفاع اشارة نسبة الضوضاء (Peak Signal to Noise Ratio (PSNR) لقياس جودة الصورة.

المصطلحات الرئيسة للبحث/ معالجة صورية، تقطيع صوري، خوارزميات عنقدة، (K-Means).

Journal of Economics and Administrative Sciences
2019; Vol. 25, No.111
Pages: 466- 484

* بحث مستل من رسالة ماجستير

1 – المقدمة وهدف البحث:

للصور دور مهم في التطبيقات، اذ تعد واحدة من اهم وسائل نقل المعلومات واستخلاصها والتي يمكن استخدامها في مهام اخرى ضرورية، واولى الخطوات باتجاه فهم الصور هو بتقطيعها واستخلاص الاهداف المتعددة منها، ويحظى التقطيع الصوري بأهمية كبيرة في مجالات علمية وصناعية متنوعة، اذ دخل في العديد من التطبيقات مثل الطبية والمخبرية ومقاييس الجودة والرؤية الحاسوبية وغيرها الكثير، ويهدف التقطيع الصوري الى تجزئة الصور المدروسة الى مناطق متعددة حسب اهداف معينة، والاستشعار عن بعد هو احد اهم الحقول المرتبطة بالتطبيقات الصورية والمؤثرة على التطبيقات الحياتية للانسان، اذ بإمكان هذا المفهوم ان يجهزنا بصور متعددة الاطياف تساعدنا على دراسة ابعاد مختلفة للاهداف المدروسة والتي لا يمكن للعين البشرية ان تلاحظها، وفي حقل الاستشعار عن بعد يفهم التقطيع الصوري على انه البحث عن المناطق المتجانسة في صورة معينة ومن ثم تصنيف هذه المناطق، وايضاً يعني تجزئة الصورة الى مناطق مفهومة بالاعتمادة على معيار تجانس معين، والتصنيف يشير الى عملية تجميع المتغيرات في مجاميع متجانسة وفق خصائص مشتركة بين هذه المتغيرات، وللتجميع او العنقدة اساليب وتقنيات عديدة منها اسلوب (K-Means) للعنقدة والذي يعد واحداً من اكثر اساليب العنقدة شيوعاً وهو من الاساليب غير المراقبة او غير المرٌشدة (Unsupervised Technique) اي تلك الاساليب التي لاتحتاج الى معلومات مسبقة يمكن اعتمادها لغرض التصنيف، وتتميز بجودتها وكفائتها في عنقدة المتغيرات، ويعد هذا الاسلوب من طرائق التحليل العنقودي غير الهرمية، وبهدف الحصول على مناطق الاستفادة يتم تقطيع الصورة الى عدة مناطق لغرض تصنيفها وبيان اهمية المنطقة المدروسة تم توضيف اسلوب (K-Means) الذي يعمد الى تجميع العناصر الصورية (Pixels) للصور المدروسة في (K) من العناقيد المتكونة من عناصر صورية تشترك في صفات محددة تختلف فيها عن العناصر الصورية للعناقيد الاخرى، لذا يهدف البحث الى تطبيق خوارزمية (K-Means) للعنقدة لغرض لتقطيع الصوري (شطر ودمج المناطق)، ومحاولة تسليط الضوء على اهمية هذا الاسلوب عند استخدامه في المعالجات الصورية من خلال تقييم اداء خوارزميته عبر تطبيقها على صورة رقمية متعددة اطياف مأخوذ لموقع في غرب العراق اذ اعتمد هذا الموقع في الدراسة نظرا للتنوع البيئي الذي يتميز به ومن ثم اختلاف في مستويات اضاءة او شدة العناصر الصورية المكونة للصورة، فهو يحوي مسطحات مائية ومناطق حضرية ومناطق نباتية واخرى مساحات خالية، ومن اهم الدراسات السابقة المستخدمة لهذه الخوارزمية، في عام (2013) عرض (Patel, P.) واخرين [XII]، أسلوب عنقدة جديد للتقطيع الصوري بأستخدام خوارزمية (K-Means) للعنقدة لغرض اكتشاف الأورام في التطبيقات الطبية والتي من الممكن ايضاً ان تطبق على صور عامة او صور محددة (صور طبية ومكروسكوبية ومجهرية) التقطت باستخدام ماسحات ضوئية وغيرها، والخوارزمية التي اعتمدها الباحثين في بحثهم توظف مفاهيم الضبابية والانتمائية لغرض إعطاء نتائج عنقدة أفضل وأكثر تكيفاً مقارنة مع خوارزميات عنقدة عديدة تقليدية، اما في عام (2015) استخدم كلاً من (Dhanachandra, N.) واخرين [VI]، خوارزمية (K-mens) للعنقدة وايضاً اقترحوا اسلوب عنقدة جديد اعتمدوه لغرض توليد المراكز الأولية المستخدمة للعنقدة ومن ثم استخدام اسلوب (K-mens) لتجزئة الصور بالاعتماد على هذه المراكز المولدة وأخيراً طبقوا تقنية التصفية الوسطية (Median Filtering) على الصور المجزئة لإزالة أية منطقة غير مرغوب بها من الصور، وفي عام (2017) اقترح (Hassan, R.) واخرين [VIII]، استخدام خوارزمية (K-Means) الاوتوماتيكية للعنقدة حيث تعتمد هذه الخوارزمية على التوليد بشكل تلقائي لمراكز العناقيد المطلوبة ومن ثم تنفيذ خطوات الخوارزمية وبهذا لا يتدخل الباحث في تحديد بعض خطوات الخوارزمية كفرض قيم أولية لمراكز العناقيد.

2 – **المعالجة الصورية** (Image Processing):

تشير المعالجة الصورية الى الحقل الذي عني بشكل اساسي باستخلاص المعلومات المفيدة من الصور من خلال اجراء عمليات حسابية او رياضية وتطبيق خوارزميات حاسوبية لتنفيذ طرائق المعالجة على الصور الرقمية والتي تستخلص هذه المعلومات مثل التحسين والتحويل والتقطيع وغيرها [VII, pp.3]، ويعد هذا الحقل مكتمل الجوانب لكنه مستمر بالتطور بسبب امكانيات الاجهزة الخارقة من سرعة وسعة تخزين والتي ساعدت بدورها على فتح افاق واسعة في عمليات المعالجة الصورية من خلال تطبيق خوارزميات وتقنيات متنوعة واحياناً معقدة كانت تعتبر في فترات معينة من المهام التي يصعب تنفيذها بامكانيات الاجهزة المتوفرة حينها، وهذا التطور في خوارزميات المعالجة الصورية مكننا من دراسة انواع مختلفة من الصور الرقمية والتي تتميز بتباين كبير بالقنوات اللونية المكونة لها، وللتعرف اكثر على مفهوم الصورة الرقمية وخصائصها نوجز المفاهيم الاساسية الاتية.

3 – **الصورة** (Image):

هي الوسيلة الاكثر نفعاً وشيوعاً في نقل او ارسال البيانات فهي تُغني عن الاف الكلمات وهي تنقل بإيجاز معلومات عن المواقع والاحجام والعلاقات المتبادلة بين الكائنات، وان حوالي (75%) من المعلومات التي يستقبلها البشر تكون على هيئة معلومات صورية [XI, pp.81]، وتصنف الصورة على نوعين اساسين هما الصورة التناظرية (Analog Image) والتي هي صورة ثنائية الابعاد كالصور الفوتوغرافية (Photographic Images) والصور التلفزيونية (Television Images)، والنوع الثاني للصورة هو الصورة الرقمية (Digital Image) والتي تعد تمثيلاً للصورة الثنائية على الحاسوب بواسطة الصفر والواحد (1,0)، والصورة الرقمية هي عبارة عن مصفوفة (Matrix) تحوي على (M) من الصفوف و(N) من الاعمدة [VII, pp.13] والتي يمكن تمثيلها كالاتي:

$$F(x,y) = \begin{pmatrix} f(1,1) & f(1,2) & \dots & f(1,N) \\ f(2,1) & f(2,2) & \dots & f(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ f(M,1) & f(M,2) & \dots & f(M,N) \end{pmatrix} \quad (1)$$

اذ تتكون هذه المصفوفة F(x,y) من العناصر الصورية (Pixels) وهي اصغر وحدة في الصورة وان دليلي الصفوف والاعمدة لهذه المصفوفة يحددها مكان العنصر (x,y) في الصورة الرقمية (F)، وان لكل عنصر من عناصر الصورة الرقمية قيمة عددية تمثل الدرجة اللونية (Color Scale) او السطوع (Brightness) او التدرج للسوية الرمادية (Gray Level) كما تعرف في حقل المعالجة الصورية والتي يمكن توضيحها بأنها مستوى التفاوت في شدة اضاءة العناصر الصورية والتي تتراوح بين (0 – 255) اذ يمثل الصفر اللون الاسود القاتم و (255) اللون الابيض الناصع [XIV, pp.12]، وبالتالي تكون الصورة الرقمية عبارة عن مصفوفة متكونة من العناصر الصورية المتفاوتة في مستويات اضاءتها والمنتشرة داخل الصفوف والاعمدة، وللصور الرقمية عدة انواع تعتمد على العناصر التي تأخذها وهذه وهي:

1.3 – الصورة الثنائية (Binary Image):

وهي من ابسط انواع الصور وتسمى ايضاً بالصورة احادية اللون (Monochrome) ويمثل هذا النوع من الصور باللونين الأسود والابيض، اي ان كل عنصر صوري يأخذ أحد القيمتين أما (0) أو (1) (اما اسود قاتم او ابيض ناصع) ويستخدم هذا النوع في صور الفاكس وغيرها [I, pp.157].

2.3 – الصورة الرمادية (Gray Scale Image):

هي صورة احادية اللون (Monochrome) لكن تأخذ تدرجات السوية الرمادية، اي ان العناصر الصورية لهذا النوع من الصور يأخذ احد اللونين اما ابيض او اسود لكن بتدرجات متفاوتة في الشدة اللونية اي تأخذ المدى (0 – 255) بدلا من (0 و1) [I, pp.157].

3.3 – الصورة الملونة (Color Image):

هي الصورة الرقمية التي تدعم الالوان عن طريق تخصيص ثلاث خانات في كل عنصر لتحديد شدة الالوان الثلاثة الاساسية (الاحمر والازرق والاخضر)، وايضاً يمكن ان تتكون الصورة الملونة من دمج ثلاث صورة احادية اللون كل صورة تتألف من احد الالوان الاساسية ومن ثم تعرض كصورة ملونة (RGB) [I, pp.158].

4.3 – الصورة متعددة الاطياف (Multi-Spectral Image):

الصورة متعددة الاطياف هي الصورة التي تحوي على معلومات تقع خارج مدى الابصار البشري وهذه المعلومات تكون على هيئة اطياف ضوئية كالأشعة تحت الحمراء (Inferred) والاشعة السينية (X-ray) او موجة الرادار (Radar Wave)، وهذه الصور لا تلتقط بالاجهزة الشائعة كالكاميرات بل تؤخذ عبر الاجهزة او التطبيقات التي تعتمد على مبدأ الاستشعار عن بعد وذلك لكون المعلومات المعروضة لا تكون مرئية بشكل مباشر بالنظام البشري، مع ذلك فان المعلومات تعرض بالصورة او بالهيئة المرئية وذلك بعد رسم او تصوير او تخطيط الحزم الطيفية المختلفة وتحويلها الى مكونات ملونة (RGB) [I, pp.158].

4 – **الاستشعار عن بعد** (Remote Sensing):

يعرف مفهوم الاستشعار عن بعد بأنه عملية جمع البيانات المتعلقة بالهدف المنشود من دون اي اتصال فيزيائي مع الهدف [II, pp.23]، مثل اخذ الصور لسطح الارض من خلال الاقمار الصناعية او الطائرات، وفي الواقع يعتمد الانسان على هذا المبدأ اذ يستعمله من خلال حواس سمعه وبصره التي تجمع المعلومات من محيطه الاجتماعي من دون الاتصال الفيزيائي به، اذ يعمد هذا المفهوم على رصد وتسجيل الطاقة الاشعاعية المنبعثة او المنعكسة عن سطح الارض على شكل موجات طولية مختلفة من الطيف الكهرومغناطيس، اما المستشعر (Sensor) فهو عبارة عن الاداة التي تستقبل الانبعاث الكهرومغناطيسي وتحوله الى اشارة (Signal) يمكن ان تسجل وتعرض اما على هيئة ارقام او على هيئة صورة [XIII, pp.3]، ويعد برنامج الرصد الفضائي (Landsat) احد اهم تطبيقات هذا المفهوم والذي يجهزنا بصور متعددة اطياف، وهو مكون من عدد من الاقمار الصناعية التي تحمل اجهزة المسح او الرصد و(Landsat Thematic Mapper) هو احد هذه الماسحات والمركب على القمرين (Landsat 4 & 5) حيث يحوي على سبعة حزم طيفية (Spectral Bands) تغطي مناطق الطيف المرئية (الزرقاء والخضراء والحمراء) والاشعة تحت الحمراء القريبة والمتوسطة والبعيدة (Near, Mid & Far Inferred) والحرارية (Thermal)، جميعها تعمل على رصد سطح الارض وتسجيل الانبعاثات المنعكسة منه لتكوين صورة متعددة اطياف (Multispectral Image) مكونة من ثلاث من اي من الاطياف السبعة المذكورة اعلاه وبحسب هدف دراسة الباحث، لذا فهناك عدد كبير من الاحتمالات لتكوين صورة متعددة الاطياف والتي تمكننا من مسح وجه الارض واستكشاف اصغر او ادق التفاصيل المرغوبة.

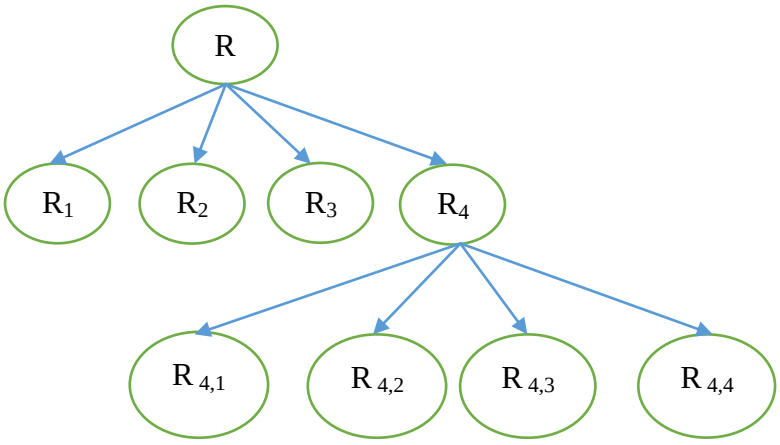
5 – **التقطيع الصوري** (Image Segmentation):

التجزئة او التقطيع الصوري هو احد اهم طرائق المعالجة الصورية وهي عبارة عن عملية تجزئة صورة معينة الى عدة اجزاء اكثر نفعاً تسمى القطع (Segments) تتشابه ببعض الخصائص او الصفات، والهدف الرئيس للتجزئة هو الحصول على صورة ذات عرض ابسط واسهل في التحليل اذ تعمد على تحديد المناطق المرغوب دراستها وفصلها عن المناطق غير المهمة عبر استخدام احد اساليب التقطيع العديدة المختلفة، اذ يشتمل التقطيع الصوري على العديد من الطرائق اذ ان هناك طرائق معتمدة على الحواف (Edge Based Methods) هذا النوع من الطرائق يعتمد على تحديد الاطر والحواف للاهداف المدروسة بغية فصلها عن بقية ملامح الصورة، وطرائق معتمدة على المناطق (Region Based Segmentation Methods) وتعمد هذه الطرائق على تقطيع الصور الى مناطق مختلفة تشترك بخصائص محددة وهي على نوعين طرائق لدمج المناطق (Region Merging) واخرى لشطر المناطق (Region Splitting)، واخيراً طرائق العتبة (Thresholding Segmentation) والتي يتمثل مبدأ عملها على تحويل الصور المدروسة الى صورة رقمية ثنائية ذات منطقتين امامية وخلفية [X, pp.811].

6 – **التقطيع باستخدام دمج وشطر المناطق** (Region Merging & Region Splitting Segmentation):

يهدف التقطيع الى تبسيط عرض الصورة المجزئة بهدف ان تكون اكثر قابلية للتفسير، والتقطيع الصوري يمكن ان ينفذ بأساليب وتقنيات عديدة منها من ينتمي الى تقنيات المعالجة الصورية المختلفة مثل طرائق معتمدة على الحواف واخرى معتمدة على المناطق وايضاً العنقدة والعتبة المذكورة في الفقرة السابقة، ومنها من يعتمد على اساليب احصائية قادرة على تنفيذ مهمة التقطيع الصوري مثل التحليل العنقودي (Cluster Analysis) والارتباط القويم (Canonical Correlation) وغيرها من الاساليب والتي يمكن ان تصنف على انها اساليب لشطر ودمج المناطق.

شطر ودمج المناطق هي عملية معتمدة على منهجية التقسيم، اذ ان مفهومها الاساسي مبني على اساس الشجرة الرباعية والذي يعني ان الصورة كمرحلة اولى تقسم الى اربعة مناطق لو كانت احداها غير متجانسة تقسم الى اربعة مناطق اخرى وهكذا وصولاً الى مرحلة عدم قابلة التقسيم وثم تأتي بعدها مرحلة الدمج وذلك بدمج المناطق المتجانسة او الاكثر تقارباً في الصفات باستخدام قاعدة دمج مناسبة لغرض اعطاء نتائج التقطيع النهائية للصورة المدروسة والتي تتمثل بصورة قد تكونت من مناطق مقطعة من الصورة الاصلية واخرى مدمجة مع مناطق اخرى لتجانسهما معاً [III, pp.307]، والهدف الرئيس لشطر ودمج المناطق هو الوصول لحالة التجانس في الصورة [XV, pp.5]، والشكل الاتي يوضح مراحل التقسيم بحسب آلية شطر ودمج المناطق:



(A)

R_1	R_2	
R_3	$R_{4,1}$	$R_{4,2}$
	$R_{4,3}$	$R_{4,4}$

(B)

الشكل (1) يوضح آلية عمل تقنيات الشطر والدمج اذ (A) توضح هيكلية الشجرة الرباعية حيث (R) هي منطقة الصورة باكملها، و(B) توضح التقطيع المقابل للشجرة الرباعية بالنسبة للصورة [15, pp.6] الاساليب الاحصائية هي اساليب تنتمي الى تقنيات الشطر والدمج، فهي تسعى الى استخلاص مناطق متجانسة متكونة من العناصر الصورية والمتجمعة وفق معايير معينة كان تكون شدة اضاءة العناصر الصورية او الحجوم او المسافات او خصائص احصائية اخرى تقاس من الصورة كدرجات الارتفاع والانخفاظ في المدرجات التكرارية، والاساليب الاحصائية كالتحليل العنقودي تسعى بدورها الى تكوين مناطق عناقيد صورية تتكون من عناصر الصورة بالاعتماد على درجة التشابه او تتكون من العناصر الصورية المتجاورية وعنصر اساسي محدد يعتمد كمركز تتعنقد حولة بقية العناصر الصورية، لذا استخدمت تقنية (K-Means) للعنقدة لغرض التقطيع الصوري كأحد اساليب التحليل الاحصائي متعدد المتغيرات للتقطيع الصوري.

7 - **التقطيع الصوري باستخدام العنقدة الاحصائية** (Image Segmentation By Clustering):

يعد التقطيع الصوري مكوناً حاسماً او حساس في نظام المعالجة الصورية بسبب كون الخطأ في هذه المرحلة يؤثر على مهام استخلاص الصفات او التصنيف ومن ثم التفسير، اذ تعني معالجة الصور الرقمية او ببساطة المعالجة الصورية تحويل الصور الى بيانات قابلة للاستخدام بمساعدة الحواسيب الرقمية حيث هناك سببان رئيسيان لمعالجة الصور باستخدام الحواسيب وهما لتحسين جودة الصور وتسهيل التفسير البشري والتصنيف التلقائي للاجسام او المفردات المختلفة التي تظهر في الصور، ويرتبط التقطيع الصوري بمشاكل العنقدة اذ يهدف التحليل العنقودي الى اتمته النشاط او الطبيعة البشرية في استخدام التصنيف لتجميع اصناف او مجاميع تتكون من مفردات تشترك في صفات محددة بغض النظر عن اصولها او طبيعتها، هذه المفردات يمكن ان تكون المرضى في المستشفيات او انواع مختلفة من منتجات استهلاكية، او الطلاب في الجامعات، او فصائل مختلفة من النباتات، او الكتب في المكتبة، او العناصر الصورية في الصور الرقمية وغيرها.

وتعد العنقدة من اساليب التصنيف ايضاً وهي اساليب تنتمي الى التحليل الاحصائي متعدد المتغيرات ويمكن استخدامها للتقطيع الصوري اذ يعد التصنيف احد اهداف التحليل الاحصائي متعدد المتغيرات، وتصنف اساليب التصنيف في التحليل الاحصائي متعدد المتغيرات الى صنفين اساسيين هما التصنيف المراقب او المرّشد (Supervised Classification) والتصنيف غير المرّشد (Unsupervised Classification)، ويقصد بالأساليب غير المرّشدة هي تلك الاساليب التي لا تحتاج الى معلومات معينة مسبقة يمكن اعتمادها لغرض التصنيف او التجميع (طريقة شطر ودمج المناطق بالتحديد) وعلى النقيض منها الطراق المرشدة والتي تعتمد معلومات مسبقة لاغراض التصنيف، ويتشابه التقطيع الصوري بنفس العلاقة مع التصنيف الصوري ومنطلقاً من هذه التشابه تم توظيف خوارزمية (K-Means) للتقطيع الصوري علماً ان التقطيع الصوري هو اكثر تعقيداً من التصنيف، اذ في مشاكل التصنيف يتم تحديد الاصناف مقدماً بينما في مشاكل التقطيع الصوري يكون عدد الاصناف غير معلوم وخوارزمية التقطيع يجب ان تتضمن بعض الاوساط الحسابية لتحديد العدد الفعلي للاصناف من البيانات، وايضاً في مشاكل التصنيف الاهداف التي يسعى الى تصنيفها هي صور جزئية، اما في مشاكل العنقدة الاهداف التي تصنف هي العناصر الصورية المفردة (Pixels).

8 – **خوارزمية التقطيع الصوري باستخدام طريقة** (K-Means) **للعنقدة**:

تم توظيف تقنية (K-Means) في التقطيع الصوري باستخدام الشطر والدمج، وهي احد اساليب التحليل الاحصائي متعدد المتغيرات والمنتمية الى طرائق العنقدة للتحليل العنقودي قام بتنظيمها العالم الامريكي (James B. McQueen) عام (1967) وتعتبر تقنية عنقدة غير مرّشدة (Unsupervised Clustering) [IV, pp.1]، وهو اسلوب اصبح مؤخراً واحداً من اوسع طرائق العنقدة انتشاراً فهو يهدف الى تجزئة مجموعة البيانات المدروسة الى (K) من مناطق العناقيد المتجانسة والمنفصلة اذ تتكون كل منطقة من عناصر تشترك بصفات معينة وتختلف مع صفات العناصر المكونة لمناطق العناقيد الاخرى، وتعتمد المسافات الفاصلة بين المفردات ومراكز مناطق العناقيد كمقياس لهذه الصفات المشتركة، ويمكن قياس هذه المسافات بواسطة مقاييس عدة لقياس المسافات تستخدم في علم الرياضيات مثل مقياس منهاتن (Manhattan) او مقياس مربع كاي (Chi-Square) او مقياس مهلنبيس (Mahalanobis) وغيرها من المقاييس، ويعد مقياس المسافة الاقليدية (Euclidean Distance) هو الاكثر شيوعاً لقياس المسافات بين العناصر وهو عبارة عن خط مسافة مستقيم مباشر بين نقطتين يقيس المسافة بينهما في الفضاء الاقليدي [IX, pp.673]، ويمكن حساب المسافة الاقليدية بحسب الصيغة الاتية:

$$d_{ij} = \sqrt{\sum_{r=1}^{p} (X_{ir} - X_{jr})^2} \quad , \quad \begin{cases} i, j = 1,2, \dots, N \\ r = 1,2, \dots, p \end{cases} \qquad (2)$$

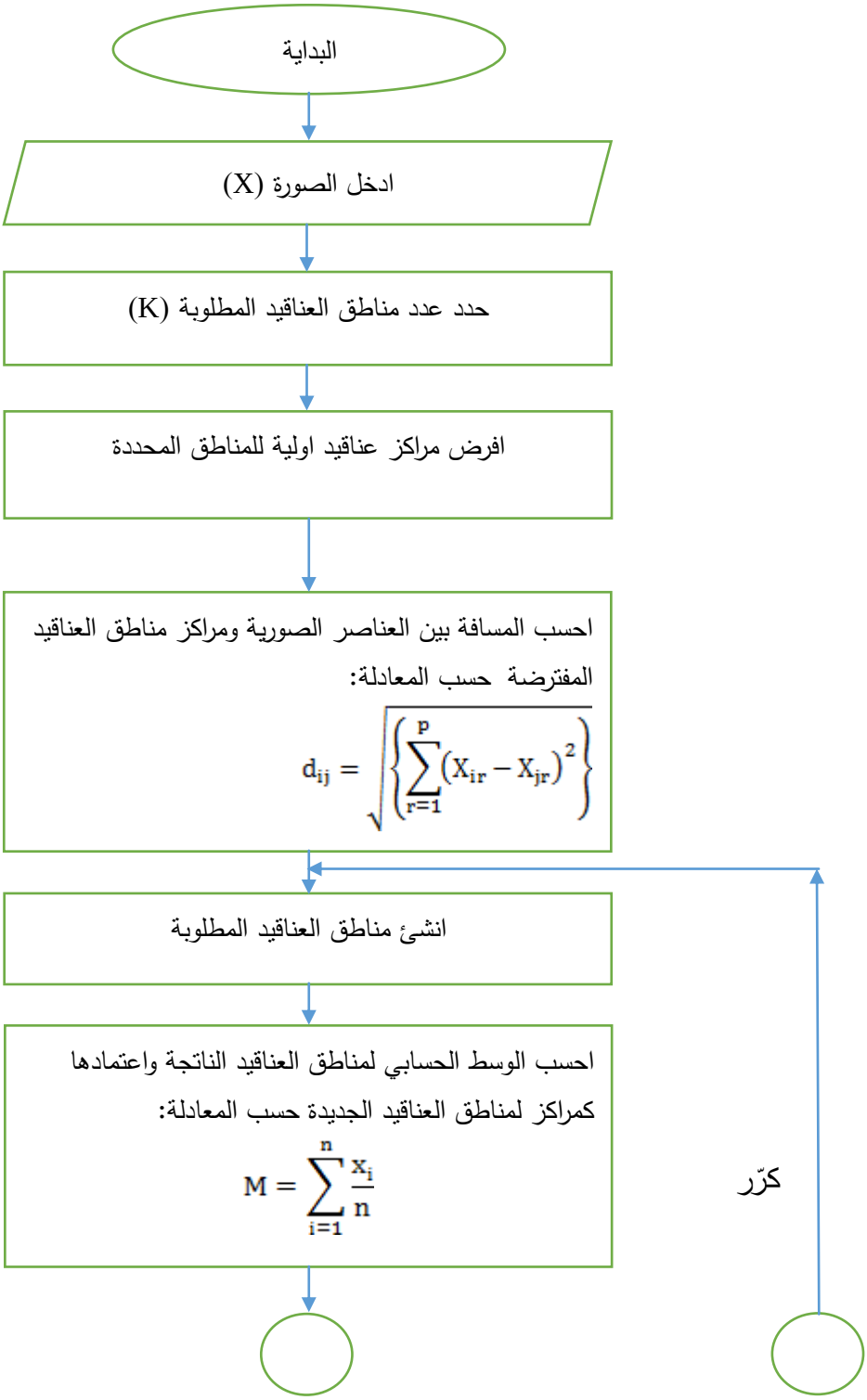
حيث (r) تمثل عدد المتجهات او عدد الابعاد في الفضاء المدروس و(i,j) هي عدد المفردات في كل متجه، وبعد حساب المسافات تتشكل مناطق العناقيد من العناصر التي تفصلها اقصر المسافات مع المراكز المفترضة، لذا فان آلية عمل خوارزمية تقنية (K-Means) تتمثل بسلسلة من الخطوات المتتالية والموضحة كالاتي:

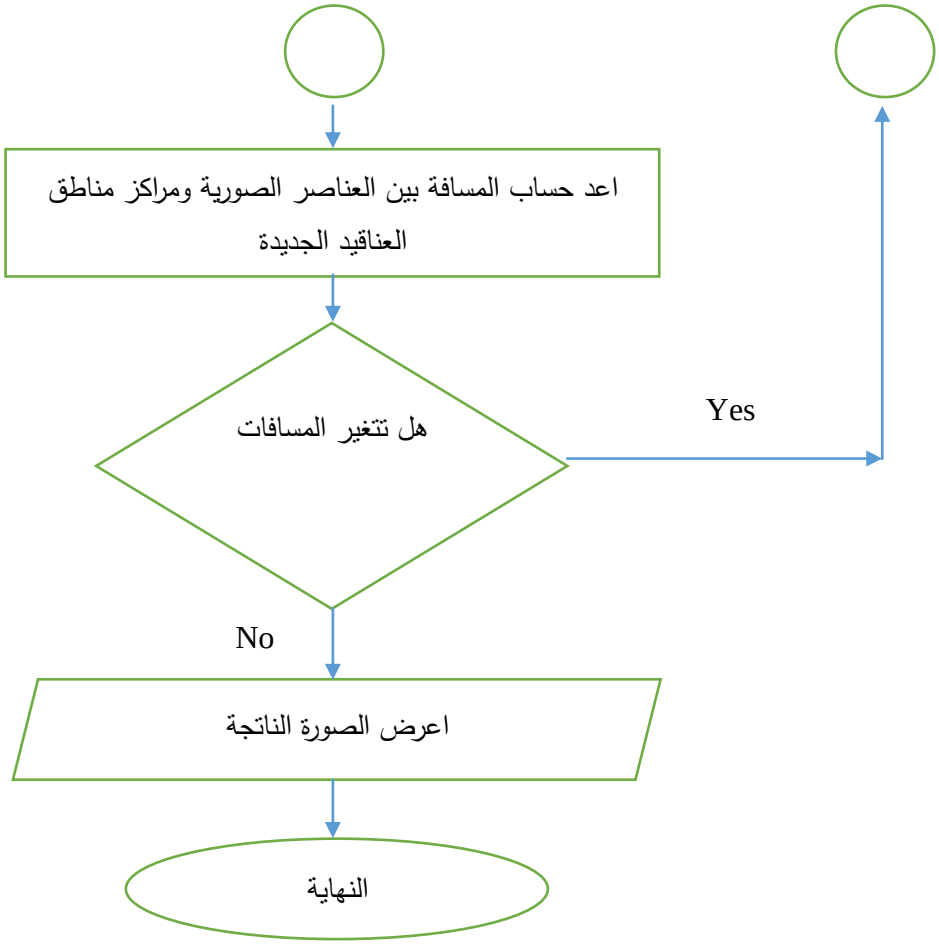
الخطوة الاولى: ندخل الصورة المعتمدة.
الخطوة الثانية: نحدد عدد مناطق العناقيد (K) المطلوب التقطيع على اساسها.
الخطوة الثالثة: نحدد مراكز مناطق العناقيد المرغوبة وذلك بفرض قيمة اولية لهذه المراكز والتي عددها مساوي الى (K).
الخطوة الرابعة: نحسب المسافة بين العناصر الصورية في الصورة المدروسة وبين المراكز المختارة.
الخطوة الخامسة: نحسب مراكز المناطق الجديدة من خلال حساب الاوساط الحسابية لكل منطقة مستخرجة واعتماد قيمها كمراكز جديدة وحسب الصيغة:

$$M = \sum_{i=1}^{n} \frac{x_i}{n} \qquad i = 1,2, \dots, n$$

الخطوة السادسة: نكرر الخطوتين (الرابعة) و(الخامسة) لحين الوصول الى مناطق عناقيد لا تتغير مفرداتها كثيراً.
الخطوة السابعة: عرض الصورة الناتجة بعد انتهاء عملية العنقدة.

وهكذا وبعد الحصول على مناطق العناقيد الجديدة تستمر العملية بالتكرار من الخطوة الثانية الى الرابعة الى ان تصل الى مناطق عناقيد لا تتغير العناصر المكونة لها بشكل لا يؤثر في خصائص المنطقة، وفيما يأتي مخطط يوضح خطوات تنفيذ خوارزمية (K-Means) للعنقدة.


البداية
ادخل الصورة (X)
حدد عدد مناطق العناقيد المطلوبة (K)
افرض مراكز عناقيد اولية للمناطق المحددة
احسب المسافة بين العناصر الصورية ومراكز مناطق العناقيد المفترضة حسب المعادلة:
$d_{ij} = \sqrt{\left\{\sum_{r=1}^{p}(X_{ir} - X_{jr})^2\right\}}$
انشئ مناطق العناقيد المطلوبة
احسب الوسط الحسابي لمناطق العناقيد الناتجة واعتمادها كمراكز لمناطق العناقيد الجديدة حسب المعادلة:
$M = \sum_{i=1}^{n} \frac{x_i}{n}$
كرّر



المخطط الانسيابي (1) يوضح خطوات خوارزمية (K-Means) للتقطيع
*** خطوات ومخطط الخوارزمية من اعداد الباحث**

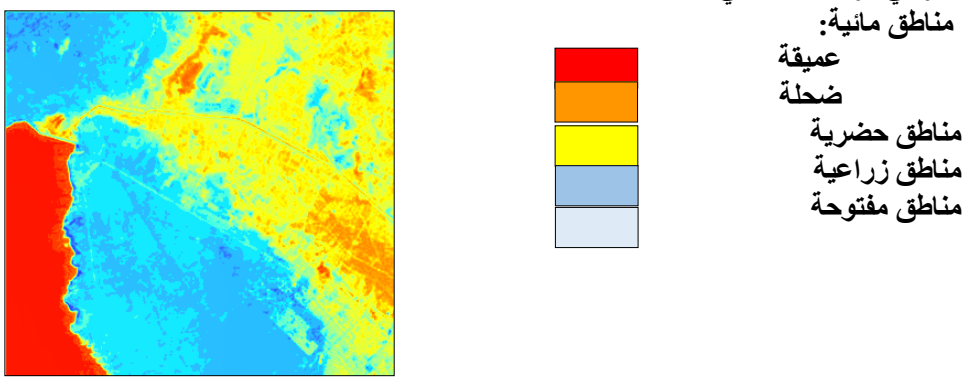
وبعد الحصول على الصورة الناتجة من تنفيذ تقنية التقطيع المقترحة وبهدف معرفة مدى جودة هذه الصورة، اعتمد مقياس احصائي يُعنى بقياس مدى جودة الصور الناتجة بعد تطبيق اي من اساليب المعالجة الصورية عليها مثل التحسين او التحويل او التقطيع، وهذا المعيار هو مقياس ارتفاع اشارة نسبة الضوضاء (Peak Signal to Noise Ratio (PSNR)) وهو عبارة عن الفرق بين العناصر الصورية بين صورتين [V, pp:10]، والذي يحسب وفق الصيغة ادناه:

$$PSNR(X,Y) = 10 * \log_{10} \frac{(MAX^2_{pixels})}{(MSE)} \quad (3)$$

حيث (X) الصورة الاصلية و(Y) هي الصورة الناتجة بعد اجراء عمليات المعالجة عليها، و(MAX) هي اعلى قيمة اضاءة للعناصر الصورية في الصورة المدخلة وهي عادة (255)، و(MSE) هو متوسط مربعات الخطأ بين الصورتين المدخلة والناتجة والذي يحسب وفق الصيغة الاتية:

$$MSE = \frac{\left(\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left(X_{(i,j)} - Y_{(i,j)}\right)^2\right)}{M * N} \quad (4)$$

حيث (M) و(N) هي العناصر الصورية في الصفوف والاعمدة، ومن ثم يسعى هذا المعيار الى قياس مدى انخفاض او ارتفاع نسبة التشويه او التغبيش (blur) للصور المعالجة اذ كلما ارتفع قيمة (PSNR) عن الصفر دل على زيادة في جودة الصورة الناتجة، وتقاس قيم هذا المعيار بوحدة الديسيبيل ((Decibel (DB) وهي وحدة لوغاريتمية تعطي النسبة بين كميتين فيزيائيتين مثل شدة الصوت وشدة الضوء وغيرها، وبالتالي ولتوضيح آلية عمل تقنية (K-Means) للتقطيع الصوري نفذت خوارزمية هذا الاسلوب باستخدام لغة البرمجة (Matlab R2017a) على صورة فضائية متعددة اطياف ملتقطة للمنطقة المدروسة وتقع غرب العراق قرب مدينة الرمادي (flight path 169 & row 37°) ومجهزة من قبل مركز أبحاث الفضاء العراقي، وهي تتصف بتنوعها البيئي ومن ثم تعد مثالاً جيداً للدراسة بسبب تباين اضاءة العناصر الموجودة في المشهد، وهي موضحة كما في ادناه:

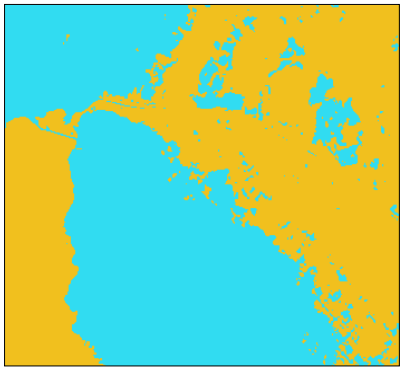


الشكل (2) يوضح الصورة المدروسة من خلال صورة ملتقطة متعددة اطياف وفئات وتقسيمات هذا المشهد

وكما هو واضح نلاحظ التباين بين مكونات المشهد اذ يتكون من مناطق مائية ضحلة تتمثل بالبرك والمستنقعات المتشكلة بسبب المبازل والقنوات الاروائية الزراعية، واخرى عميقة، وايضاً مناطق غطاء نباتي يتمثل بالمحاصيل الزراعية للمزارع المنتشرة حول القنوات الاروائية واخرى مناطق مفتوحة خالية، واخيراً الغطاء الحضري المتمثل بالمناطق السكنية المنتشرة قرب المناطق الزراعية.

9 – النتائج والتفسير:

اسلوب عنقدة (K-Means) يتميز بقدرتة على تحديد او التحكم بعدد مراكز مناطق العناقيد المرغوب اعتمادها لغرض تكوين المناطق المنشودة لذا استهدفنا استخراج او تكوين صور في حالات عدد مناطق عناقيد مختلفة وهي (5,4,3,2)، والنتائج هي نتيجة تكرار خوارزمية (K-Means) العمليات الحسابية الى الوصول الى حالة الاستقرار في النتائج، وجاءت نتيجة التقطيع للحالة الاولى المعتمدة على اساس عنقودين جاءت النتائج كالاتي:



المنطقة الاولى (المناطق الزراعية والمناطق المائية)
المنطقة الثانية (المناطق السكانية والمناطق المفتوحة)

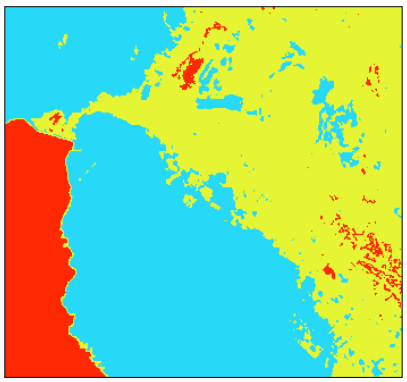
الشكل (3) يوضح الصورة المستحصلة نتيجة التقطيع على اساس منطقتين

والشكل المذكور آنفاً يوضح نتيجة التقطيع لمنطقتين باستخدام خوارزمية (K-Means)، اذ ان العناصر الصورية تجمعت في منطقتين اساسيتين تضمنت المنطقة الاولى المناطق الزراعية والمائية معاً، بينما تضمنت المنطقة الثانية المناطق السكانية والمفتوحة، اما الجدول الاتي فيعرض مراكز العناقيد للمنطقتين الموجودتين في الصورة والتي تجمعت حولها بقية العناصر الصورية، وايضاً متوسط مربعات الاخطاء لكل عنقود:

الجدول (1) يوضح مراكز العناقيد ومتوسطات مربعات الاخطاء لهذه العناقيد للصورة ذات المنطقتين

تفاصيل العناقيد	عدد العناقيد		
	1	2	
مراكز العناقيد	96.32	191.86	الصورة الناتجة ذات العنقودين
متوسط مربعات الاخطاء للعناقيد	0.5306	0.4693	

من الجدول يتبين ان مركز العنقود الاول يساوي (96.32) وهو يعكس سبب كون منطقة هذا العنقود هي الاكبر بسبب النسبة المرتفعة للعناصر الصورية التي شدتها قريبة من هذا المقدار، في حين العنقود الثاني كان مركزة هو (191.86) وهو الوسط الحسابي لهذا العنقود، وايضاً حسبت قيمة متوسط مربعات الاخطاء لكل عنقود ناتج حيث كانت قيمتها للعنقود الاول (0.5306) وللعنقود الثاني (0.4693)، اما في حالة لو اردنا تقطيع الصور الى ثلاث مناطق فأن النتيجة تكون كما في الشكل الاتي:



المنطقة الاولى (المناطق المائية والمناطق العميقة والضحلة)
المنطقة الثانية (المناطق المفتوحة والمناطق الحضرية)
المنطقة الثالثة (المناطق الزراعية)

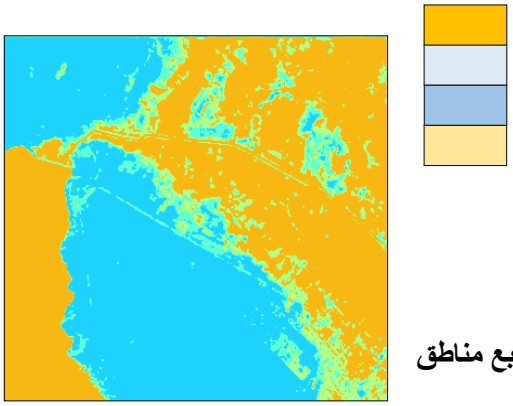
الشكل (4) يوضح الصورة المستحصلة نتيجة التقطيع على اساس ثلاث مناطق

والشكل المذكور آنفاً يوضح آلية التقسيم التي اعتمدتها الخوارزمية، اذ عمدت على فصل الصورة الى ثلاث مناطق تمثلت الاولى بالمناطق المائية العميقة والضحلة اذ كانت قيمة مركز المنطقة تساوي (91.81) والموضحة في الجدول (2) ادناه والتي تركزت حولها العناصر الصورية القريبة من هذا المقدار وبنفس الوقت حسب متوسط مربعات الخطأ لهذه المنطقة والذي يساوي (0.4336)، اما بالنسبة للمنطقة الثانية فشتملت على المناطق المفتوحة وايضاً المناطق الحضرية المتواجدة حولها والتي تكونت من العناصر الصورية المتجمعة حول مركز العنقود لها والمساوي (162.39) وكانت قيمة متوسط مربعات الخطأ لهذه المنطقة يساوي (0.1121)، اما المنطقة الثالثة والاخيرة فتمثلت بالمناطق الزراعية والتي تجمعت العناصر الصورية حول مركز هذه المنطقة والمساوي الى (218.11) وقيمة متوسط مربعات الخطأ لهذه المنطقة كانت يساوي (0.4541).

الجدول (2) يوضح مراكز العناقيد ومتوسطات مربعات الاخطاء لهذه العناقيد للصورة ذات الثلاث مناطق

عدد العناقيد			تفاصيل العناقيد	
3	2	1		
218.11	162.39	91.81	مراكز العناقيد	الصورة الناتجة ذات الثلاث عناقيد
0.4541	0.1121	0.4336	متوسط مربعات الاخطاء للعناقيد	

نلاحظ ايضاً من خلال الجدول اعلاه ان قيم متوسطات مربعات الاخطاء قد انخفظت عن حالة التقطيع لمنطقتين فالمنطقة الاولى سجلت انخفاضاً قدره (0.097) والمنطقة الثانية انخفضت بمقدار (0.3572) وهذا يوضح ان بزيادة عدد المناطق المقطع على اساسها الصورة تزداد العناقيد تجانساً وتقارباً، اما في حالة قطعنا على اساس اربع عناقيد تظهر النتائج كما يأتي:



المنطقة الاولى (المناطق المائية والمناطق الزراعية)
المنطقة الثانية (المناطق السكنية)
المنطقة الثالثة (المناطق المفتوحة والمناطق الحضرية)
المنطقة الرابعة (مناطق اضافية محيطة ببقية المناطق السابقة)

الشكل (5) يوضح الصورة المستحصلة نتيجة التقطيع على اساس اربع مناطق

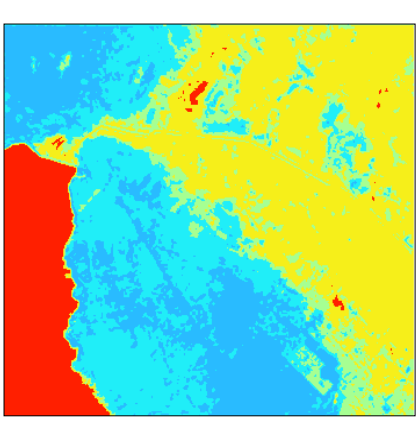
اذ بتنفيذ الخوارزمية يتم فصل الصورة الى اربعة مناطق تمثلت الاولى بالمناطق المائية وجزءاً كبيراً من المناطق الزراعية المحيطة بالقنوات والسواقي المائية وكانت قيمة مركز المنطقة تساوي (91.47) والموضحة في الجدول (3) ادناه والتي تركزت حولها العناصر الصورية القريبة من هذا المركز، في حين كانت قيمة متوسط مربعات الخطأ لهذه المنطقة تساوي (0.4187)، اما بالنسبة للمنطقة الثانية فحوت جزءاً صغيراً من المشهد، اذ عرضت هذه المنطقة طرق النقل الرابطة بين المناطق السكنية وايضاً بعضاً من المجمعات السكنية، وكانت قيمة مركز العنقود لهذه المنطقة تساوي (131.75) وقيمة متوسط مربعات الاخطاء (0.1119)، اما المنطقة الثالثة فتمثلت بالمناطق المفتوحة والحضرية والتي كانت قيمة مركز هذه المنطقة تساوي (163.52) وقيمة متوسط مربعات الخطأ لهذه المنطقة كانت تساوي (0.3635)، اما المنطقة الرابعة والاخيرة فتمثلت بمنطقة صغيرة منتشرة في المشهد المدروس وبنفس الوقت تتخلل مجمل المناطق وكأنها اطار يحدد ويفصل المناطق عن بعضها الاخر، وكانت قيمة مركز هذه المنطقة تساوي (218.11) وقيمة متوسط مربعات الخطأ لها تساوي (0.1057).

وايضاً نلاحظ ان في هذه الحالة (التقسيم على اساس اربع مناطق) هناك مناطق اختفت (المياه الضحلة) وظهرت مناطق اخرى لم تكن موجودة (مناطق اضافية محيطة ببقية المناطق)، والسبب يعود الى كون العناصر الصورية المكونة لمناطق المياه الضحلة اصبحت اقرب لمراكز مناطق اخرى بدلا من مركزها السابق، اذ وكما وضح سابقا تعمد التقنية المعتمدة على استخراج الاوساط الحسابية لكل منطقة في كل مرحلة جديدة لغرض اعتمادها كمراكز للعناقيد وتجمع العناصر الصورية عليها لغرض تكوين المناطق الجديدة، والنتائج للمناطق المتكونة موضحة في الجدول الاتي

الجدول (3) يوضح مراكز العناقيد ومتوسطات مربعات الاخطاء لهذه العناقيد للصورة ذات الاربعة مناطق

عدد العناقيد				تفاصيل العناقيد	
4	3	2	1		
218.11	163.52	131.75	91.47	مراكز العناقيد	الصورة الناتجة ذات الاربع عناقيد
0.1057	0.3635	0.1119	0.4187	متوسط مربعات الاخطاء للعناقيد	

وفي حالة التقطيع بالاعتماد على اربعة عناقيد نلاحظ ايضاً انخفاض قيمة متوسطات مربعات الاخطاء للمناطق المتكونة عن الحالة السابقة (التقطيع على اساس ثلاث مناطق) اذ انخفض قيمة الخطأ للعنقود الاول بمقدار (0.0149) في حين انخفضت بالنسبة للعنقود الثاني بمقدار (0.0002)، اما المنطقة الثالثة فشهدت انخفاضاً بمقدار (0.0906)، وارقام الانخفاض هذه جميعاً تؤكد قابلية تجانس العناقيد بزيادة المناطق المقطعة على اساسها الصورة، واخيراً ونتائج حالة التقطيع على اساس خمس مناطق عناقيد موضحة في الشكل الاتي:



المنطقة الاولى (المناطق المائية)
المنطقة الثانية (المناطق الزراعية)
المنطقة الثالثة (مناطق اضافية محيطة بالمناطق الزراعية)
المنطقة الرابعة (المناطق الحضرية)
المنطقة الخامسة (المناطق المفتوحة)

الشكل (6) يوضح الصورة المستحصلة نتيجة التقطيع على اساس خمس مناطق

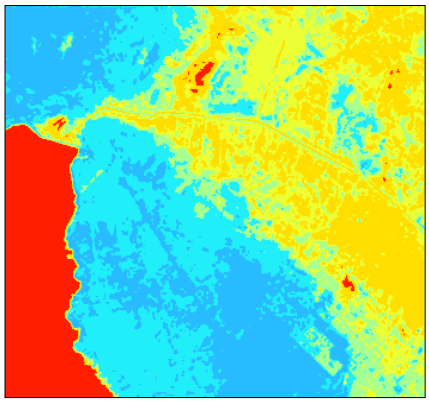
اما التقطيع على اساس خمس مناطق اعطى حالة التقطيع الاكثر تجانساً من الحالات السابقة وذلك لبقاء قيمة متوسط مربعات الاخطاء على حالها في بعض العناقيد او انخفاضها في عناقيد اخرى اكثر من الحالات السابقة المدروسة، والتقطيع لهذه الحالة ولّد خمس مناطق عناقيد، عرضت المنطقة الاولى المناطق المائية، اذ كانت قيمة مركز المنطقة تساوي (89.25)، في حين كانت قيمة متوسط مربعات الخطأ لهذه المنطقة تساوي (0.4187) وهي ذات القيمة في حالة التقسيم السابقة، اما بالنسبة للمنطقة الثانية فتمثلت بالمناطق الزراعية والتي كانت قيمة مركز عنقود منطقتها يساوي (121.11) وبلغت قيمة الخطأ لهذا العنقود (0.1119) وايضاً هي ذاتها لقيمة حالة التقطيع السابقة وهذا يدل على وصول العنقود لحالة التجانس الكاملة ومن ثم لا يوجد تغيير في العناصر الصورية المكونة لهذا العنقود وايضاً للعنقود السابق، اما المنطقة الثالثة فتمثل بمناطق زراعية ثانوية محيطة بالمناطق الاساسية المتواجدة بالعنقود او المنطقة الثانية.

اذ بلغت قيمة مركز هذا العنقود (152.73) وقيمة متوسط الخطأ لهذا العنقود بلغت (0.2499) والتي سجلت انخفاضاً عن قيمة العنقود نفسه في حالة التقسيم السابقة على اساس اربع عناقيد والتي كانت (0.3635) بالتالي تكون قد انخفضت بمقدار (0.1136)، اما المنطقة الرابعة فعرضت المناطق الحضرية والتي بلغت قيمة مركز عنقودها (172.88) وقيمة الخطأ لها (0.1314) وهي الحالة الوحيدة التي نلاحظ فيها ارتفاع الخطأ بدلاً من انخفاضها ويعود هذا السبب الى عدد العناصر الصورية وشدة اضاءة هذه العناصر المكونة للمنطقة، اما المنطقة الخامسة والاخيرة فتكونت من المناطق المفتوحة (الخالية او الجرداء) المتخللة لمشهد عموماً والمحيطة بالمناطق الحضرية لاسيما وقد بلغت قيمة مركز هذا العنقود (218.13) وقيمة متوسط مربعات الخطأ لهذا العنقود تساوي (0.0879)، والمعروضة في الجدول الاتي:

الجدول (4) يوضح مراكز العناقيد ومتوسطات مربعات الاخطاء لهذه العناقيد للصورة ذات الخمسة مناطق

	تفاصيل العناقيد	عدد العناقيد				
		1	2	3	4	5
الصورة الناتجة ذات الخمس عناقيد	مراكز العناقيد	89.25	121.11	152.73	172.88	218.13
	متوسط مربعات الاخطاء للعناقيد	0.4187	0.1119	0.2499	0.1314	0.0879

ولغرض الوقوف على مدى جودة التقطيع في حالة زيادة مناطق العناقيد المقطع على اساسها، سعينى الى تقطيع الصورة المدروسة الى (6) ومن ثم (7) مناطق عناقيد، ففي حالة التقطيع لست مناطق ظهرت الصورة الناتجة فيما يأتي:



المنطقة الاولى (المناطق المائية)
المنطقة الثانية (المناطق الزراعية كثيفة)
المنطقة الثالثة (المناطق الزراعية)
المنطقة الرابعة (مناطق اضافية محيطة بالمناطق الزراعية)
المنطقة الخامسة (المناطق الحضرية)
المنطقة السادسة (المناطق المفتوحة)

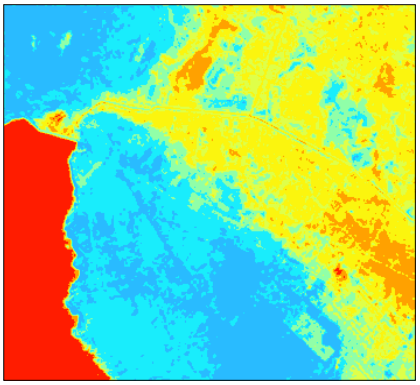
الشكل (7) يوضح الصورة المستحصلة نتيجة التقطيع على اساس ست مناطق

اذ نلاحظ ان نتيجة التقطيع على اساس ست مناطق ولدت منطقة جديدة تمثلت بالمناطق الزراعية الكثيفة التي تميزت عن بقية المناطق الزراعية، وايضا نلاحظ استمرار انخفاض متوسط مربعات الاخطاء للمناطق الجديدة الناتجة، اذ بلغ الخطأ للعنقود الاول والذي يعرض المناطق المائية بلغ (0.3924) بعد ان كان (0.4187) وايضاً بالنسبة للمنطقة الثانية والتي تعرض المناطق الزراعية الكثيفة بلغ الخطأ لها (0.1100) بعد ان كان (0.1119) وهكذا لبقة العناقيد كما موضح في الجدول (5) الاتي، في حين كان متوسط الخطأ للعنقود الجديد الناتج (العنقود السادس) مساوي الى (0.0879)، هذه الانخفاض يعود الى تقليص عدد العناصر المكونة للمناطق الجديدة نتيجة لظهور منطقة جديدة وايضا تغير قيمة مراكز العناقيد للمناطق الجديدة، اذا كما يلاحظ ان المنطقة الرابعة مثلا مركزها كان مساوي الى (152.77) في حيث في حالة التقطيع السابقة كان مساوي الى (172.88) والنتائج للمناطق الناتجة موضحة في الجدول الاتي:

الجدول (5) يوضح مراكز العناقيد ومتوسطات مربعات الاخطاء لهذه العناقيد للصورة ذات الست مناطق

	تفاصيل العناقيد	عدد العناقيد					
		1	2	3	4	5	6
الصورة الناتجة ذات الست عناقيد	مراكز العناقيد	89.15	109.93	151.59	152.77	172.88	218.13
	متوسط مربعات الاخطاء للعناقيد	0.3924	0.1100	0.2254	0.1020	0.0798	0.0879

ويلاحظ من النتائج لحالة التقطيع السابقة ان بزيادة المناطق المقطعة على اساسها الصورة تزداد المناطق تجانسا وهو ما وضح سابقاً عند التقطيع على اساس خمسة مناطق عناقيد وذلك لزيادة مسافة العناصر الصورية المكونة لكل منطقة قربٍ من المركز المكون للمنطقة، اما الحالة الاخيرة للتقطيع وهي على اساس سبع مناطق عناقيد فالصورة الناتجة لها هي كالاتي:



المنطقة الاولى (المناطق المائية)
المنطقة الثانية (المناطق الزراعية كثيفة)
المنطقة الثالثة (المناطق الزراعية)
المنطقة الرابعة (مناطق اضافية محيطة بالمناطق الزراعية)
المنطقة الخامسة (المناطق الحضرية)
المنطقة السادسة (المناطق المفتوحة)
المنطقة السابعة (مناطق محيطة بالمناطق المائية)

الشكل (8) يوضح الصورة المستحصلة نتيجة التقطيع على اساس سبع مناطق

ونتيجة التقطيع هذه كانت هي الاقرب الى المشهد المدروس قبل التقطيع فقد شملت اغلب تفاصل الصورة وايضاً شهدت استقرار بعض المناطق في عناصرها الصورية المكون لها وبالتالي استقرار متوسط مربعات الخطأ التي ظهرت في حالة التقطيع السابقة (على اساس ست مناطق عناقيد) وهي المناطق الاولى والثانية والرابعة والخامسة حيث استقرت قيم مراكزها عند (89.15) و(109.93) و(152.77) و(172.85) على التوالي، في حين تغيرت مراكز المناطق الثالثة والسادسة والتي بلغت (131.77) و(200.94) على التوالي وبلغ متوسط الخطأ لكلا المنطقتين (0.1015) و(0.0855)، في حين العنقود الجديد السابع كان المركز له (222.39) ومتوسط الخطأ له (0.925) والنتائج موضحة في الجدول الاتي:

الجدول (6) يوضح مراكز العناقيد ومتوسطات مربعات الاخطاء لهذه العناقيد للصورة ذات السبع مناطق

	تفاصيل العناقيد	عدد العناقيد						
		1	2	3	4	5	6	7
الصورة الناتجة ذات السبع عناقيد	مراكز العناقيد	89.15	109.93	131.59	152.77	172.85	200.94	222.39
	متوسط مربعات الاخطاء للعناقيد	0.3824	0.1100	0.1015	0.1020	0.0798	0.0855	0.925

واخيراً وللوقوف على مدى جودة الصور الناتجة من التقطيع حسب عدد المناطق المدروسة، تم استخدم مقياس ارتفاع اشارة نسبة الضوضاء (PSNR) المذكور سابقاً والذي اعطى القيم في الجدول الاتي مقاسة بالديسيبيل ولكل مستوى تقطيع:

الجدول (7) يوضح قيم معيار (PSNR) لقياس جودة الصور الناتجة والمكونة حسب عدد العناقيد المطلوبة

عدد العناقيد (K) للصورة الناتجة	2	3	4	5	6	7
PSNR	14.3479	18.5126	20.7254	22.8587	24.1383	25.3231

النتائج في الجدول توضح ان كلما ازدادت عدد مناطق العناقيد ازدادت جودة الصورة وهي بنفس الوقت تؤكد صحة نتائج متوسطات مربعات الاخطاء والتي كانت تشهد انخفاضاً داخل العناقيد مع كل حال تقطيع جديدة والتي تشير الى السعي الى الوصول لحالة الاستقرار والتجانس بين العناصر الصورية داخل كل منطقة الى ان وصلنا الى مرحلة التقطيع على اساس سبع مناطق والتي شهدت اغلب مراكز مناطقها استقرارا ومن ثم عدم تغير عناصرها الصورية ومن ثم تجانسها، فنلاحظ ان الصورة ذات المنطقتين بلغت مدى جودتها (14.3479 DB) وهي الاقل بين بقية الصور بينما بلغ مقدار (PSNR) لجودة الصورة المكونة من سبع مناطق (25.3231 DB) وهو الاعلى بين كل مستويات التقطيع المعتمدة وهذا يدل على انها الاكثر جودة والاقرب تجانساً من غيرها من الحالات، ومن ثم من خلال النتائج للحالات التقطيع السابقة نتوصل الى ان اسلوب عنقدة (K-Means) هو اسلوب كفوء في تخفيض متوسط مربعات الاخطاء لمناطق العناقيد المكونة في الصور الناتجة وهو الامر الذي يعتبر مؤشر على زيادة تجانس مناطق العناقيد وانخفاض مستوى التشويه او التغبيش في هذه المناطق والمكونة للصورة، وايضا ومن خلال النتائج والصور المعروضة لحالات التقطيع السابقة نلاحظ ان بزيادة المناطق المفترض التقطيع للصورة المدروسة على اساسها فأن الجودة للصور الناتجة تزداد، لكن بنفس الوقت كثرة المناطق العنقودية المتكونة تؤدي الى تشوية الادراك او الملاحظة او التفسير البشري لذا لابد من ان تكون عدد العناقيد محدد بالهدف المنشود تقطيعه، فصحيح ان النتائج اثبتت ان جودة الصور تزداد بزيادة مناطق العناقيد لكن هذا سيؤدي الى خلق عدد كبير من العناقيد المنتشرة في مناطق متقطعة من الصورة ومن ثم ستؤدي الى تشوية الملامح المطلوب عزلها لذا لابد من ان تكون مناطق العناقيد محددة لتلافي الزيادة الفائضة او حتى النقصان غير المقصود في المناطق النافعة.

10 – **الاستنتاجات**:

النتائج اظهرت مدى اهمية ان تكون عدد العناقيد (K) محددة بالقدر الذي يعطي افضل منفعة لتلافي التشويه الناتج في المعالم في حالة تحديد عدد كبير من مناطق العناقيد الذي يؤدي الى كثرة المناطق المعزولة، وايضاً لتلافي اندثار الملامح في حالة تحديد عدد اقل من مناطق العناقيد المطلوب والذي يؤدي الى دمج العديد من المناطق الضرورية.

11 – **التوصيات**:

1 – محاولة التقطيع على اساس كل قناة لونية مكونة للصورة المدروسة ومن ثم دمج القنوات الناتجة بعد التقطيع في صورة واحدة لمعرفة مدى جودة الصورة الناتجة وحسب معايير الجودة المعتمدة مثل (PSNR) وايضاً مقارنة مع التقطيع الاعتيادي.

2 – محاولة وضع معيار محدد لاعتماد العناقيد المطلوبة لغرض تقطيع الصورة وحسب الهدف المنشود لتلافي الاخطاء في زيادة او نقصان مناطق العناقيد الضرورية، واعتماد المدى بين مستويات شدة العناصر الصورية كمعيار لهذا الغرض.

12 – References:

1 - صالح، وصفي طاهر؛ أحمد، رزكار مغديد (2012) "إستخدام تحليل المركبات الرئيسية لأستخلاص المعلومات وتحسين النسجة في معالجة الصور الرقمية" مجلة جامعة كركوك للعلوم الادارية والاقتصادية، المجلد (2)، العدد (1)، ص ص (154-182).

2 - Aggarwal, Shefali (2004) "Principals of Remote Sensing" Journal of Satellite Remote Sensing and GIS Applications in Agricultural Meteorology, Vol.3, No.1182, pp.23-38.

3 - Bala, Anju & Sharma, A. K. (2017) "Split and Merge: A Region Based Image Segmentation" Journal of Emerging Research in Management &Technology, Vol.6, No.8, pp.306-309.

[IV] Burney, S. & Tariq, H. (2014) "K-Means Cluster Analysis for Image Segmentation" Journal of Computer Applications, Vol.96, No.4, pp.1-8.

4 - Dass, A. K., Shial, R. K. & Gouda, B. S. (2012) "Improvising MSN and PSNR for Finger Print Image Noised by Gaussian and Salt & Pepper" Journal of Multimedia & its Applications, Vol.4, No.4, pp.59-72.

5 - Dhanachandra, N., Manglem, K. & Chanu, Y. J. (2015) "Image Segmentation using K-means Clustering Algorithm and Subtractive Clustering Algorithm" Journal of Procedia Computer Science, Vol.4, No.54, pp.764-771.

6 - Gonzalez, R. C., Woods, R. E. & Eddins, S. L. (2004) "Digital Image Processing Using Matlab" 1th Ed., Pearson Prentice Hall, Upper Saddle River, New Jersey, USA.

7 - Hassan, R., Ema, R. R. & Islam, T. (2017) "Color Image Segmentation using Automated K-Means Clustering with RGB and HSV Color Spaces" Journal of Computer Science and Technology: Graphics & Vision, Vol.17, No.2, pp.24-33.

8 - Johnson, R. & Wichern, D. (2007) "Applied Multivariate Statistical Analysis" 6th Ed., Pearson Prentice Hall, Upper Saddle River, New Jersey, USA.

9 - Kaur, Dilpreet & Kaur, Yadwinder (2014) "Various Image Segmentation Techniques: A Review" Journal of Computer Science and Mobile Computing, Vol.3, No.5, pp.809-814.

10 - Kumar, Minakshi (2003) "Digital Image Processing" Journal of Satellite Remote Sensing and GIS Applications in Agricultural Meteorology, Vol.3, No.31, pp.81-102.

11- Patel, P. M., Shah, B. N. & Shah, V. (2013) "Image segmentation using K-mean clustering for finding tumor in medical application" Journal of Computer Trends and Technology, Vol.4, No.5, pp.1239-1242.

12 - Sanderson, R. (2010) "Introduction to Remote Sensing" New Mexico University, USA.

13 - The Math Works (2000) "Image Processing Toolbox For Use with Matlab" The Math Works, Inc.

14 - Wang, Y. Hsiang (2010) "Tutorial: Image Segmentation" Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan.

Split and Merge Regions of Satellite Images using the Non-Hierarchical Algorithm of Cluster Analysis

Abstract:

Image segmentation is one of main and necessary goals in image processing for digital image, it seeks to segment the studied images into multiple useful regions that summarizing interest regions for satellite images, which are multispectral images equipped by satellites using the concept of remote sensing, which has become an important concepts that it's applications used in most of the necessities of daily life, especially after the rapid developments in various fields of life that many of these fields have used algorithms and software techniques, these images are very necessary to enable us to study a wide range of goals in many scientific fields, in this research, the nonhierarchical cluster analysis algorithm was used as a method of image segmentation (splitting and merging regions) in order to demonstrate the importance of using statistical methods in image processing tasks, such as image segmentation, where (K-Means) technique was used to implement this task, this algorithm was applied on multispectral satellite image of a scene from western Iraq, where the results showed the flexibility of this algorithm in dealing with the disparity in the lighting of color image pixels and it's efficiency of formation clustering region that composed from groups of homogeneous pixels in there degree of illumination intensity, finally, the ability of this algorithm to give a good quality images which are measured using peak signal to noise ratio (PSNR) scale to measure image quality

Key Words: Image Processing, Image Segmentation, Clustering Algorithms, K-Means