١١ـ كيف تبنى "روبوت" حقيقى؟

الرؤيـــة الروبوتيـــة Robotic Vision

د. عـــلاء خميس
كلية هندسة البترول - جامعة قناة السويس

توفر حاسة الرؤية في الإنسان ما يزيد عن ٨٥٪ من البيانات الاستشعارية التي تساعده على التفاعل مع البيئة المحيطة به دون حدوث أضرار. ويحاول مصممو الروبوت محاكاة هذه الحاسة المعقدة من خلال تزويد الروبوت بمنظومات رؤية اصطناعية.. مما قد يفتح الباب على مصراعيه أمام كثير من التطبيقات الحديثة مثل اكتشاف الفضاء والبحار والرقابة والتجميع والفحص وضبط الجودة إلخ..

وفي الروبوتات الذكية.. تلعب منظومات الرؤية باستخدام الحاسب دورًا مهمًا جدًا.. لما توفره من قدرة للروبوت على الاستجابة والتفاعل مع البيئة المحيطة به بأسلوب مرن وذكي. ومن هنا.. يمكن تعريف رؤية الروبوت بأنها عملية استخراج وتحديد وفهم المعلومات التي توفرها الصور ثنائية أو ثلاثية الأبعاد في البيئة الحقيقية. ولعل من أهم ما يميز منظومات الرؤية الروبوتية عن منظومة رؤية الإنسان.. أن الروبوت يرى فقط ما ينفعه لأداء المهمة المكلف بها.. فروبوت التجميع على سبيل المثال يجب عليه تحديد أركان القطعة المراد تجميعها دون الحاجة لمعرفة كل صفاتها.

الرؤية البشرية Human Vision

تعتبر حاسة الرؤية في الإنسان.. عملية معقدة وتكاد تكون غير مفهومة.. وهي تتم على مرحلتين.. الأولى بواسطة العين ويتم فيها تكوين صورة باستخدام الضوء المنعكس من مؤثر خارجي Stimulus .. والثانية يقوم بها المخ ويتم فيها معالجة الصورة لفهم ماهية هذا المؤثر.

فعند دخول الضوء للعين- شكل رقم (١) - فإنه ينكسر بواسطة القرنية Cor-nea.. وهي عبارة عن نسيج قوي شفاف مقوس بشكل كروي تقوم بدور نافذة العين. وتعتبر القرنية هي عنصر التركيز الرئيسي للعين. وبعد ذلك.. تعمل الحدقة Pupil كفتحة مركزية تسمح بمرور الضوء إلى داخل العين. وتقوم القزحية Iris بتوسيع وتقليص هذه الفتحة المركزية وتساعد العدسة Lens على التركيز.. وهي عبارة عن قوس بللوري شفاف مرن محدب الوجهين ويقع خلف الحدقة حيث تسيطر عضلات علي شكل العدسة بطريقة تلقائية ليتم التركيز. ولما كانت الشبكية Retina.. عبارة عن غشاء حساس للضوء يبطن الحائط الخلفي للعين ويتكون من عدة طبقات تحتوي على ملايين الخلايا الحساسة للضوء وتسمى Rods و Cones.. فإنها تقوم بتكوين صورة مقلوبة للغرض.. ويتم تحويل الضوء إلى نبضات كهربية ترسل عن طريق العصب البصري إلى الدماغ للترجمة الفورية وفي هذه العملية.. نلاحظ أن العين تعمل كمستشعر يحول الضوء إلى نبضات كهربية يتم إرسالها عن طريق العصب البصري للمخ لمعالجتها وفهمها ليتحقق الإدراك البصري Visual Preception.. وهي عملية غير مفهومة إلى الآن. ومن التحديات الكبرى التي تواجه الروبوت.. القدرة على التعرف والتمييز بين الأغراض Object Recognition.. فالإنسان يمتلك ذاكرة ديناميكية غير معروفة أسرارها إلى الآن تمكنه من التعرف والتمييز بين الأغراض المحيطة به.. في حين أن الروبوت لا يستطيع مثلاً التمييز بين كوب وإناء موضوعين على مائدة إلا إذا تم تدريبه وبرمجته على ذلك.. الأمر الذي عادة ما يتم باستخدام منظومة رؤية روبوتية حيث يتم تخزين صورة للغرض المراد التعرف عليه في ذاكرة الروبوت. وعن طريق منظومة الرؤية.. يتم المقارنة وتحديد نوع الغرض. إلا أن مشكلة بسيطة مثل وجود مرآة على المائدة أو وجود صورة بالحجم الطبيعي للكوب أو للإناء قد تجعل منظومة الرؤية عاجزة عن التمييز بين الغرض الحقيقي والتخيلي.. ويمكن حل هذه المشكلة بتكوين صورة ثلاثية الأبعاد لقياس عمق الغرض وذلك باستخدام كاميراتين كما هو الحال في الإنسان.. حيث تتكون صورة ثلاثية الأبعاد باستخدام صورتين ثنائيتين.. إلا أن هذا الحل يتطلب وجود قدرة حاسوبية وتخزينية عالية مخصصة لمنظومة الرؤية.

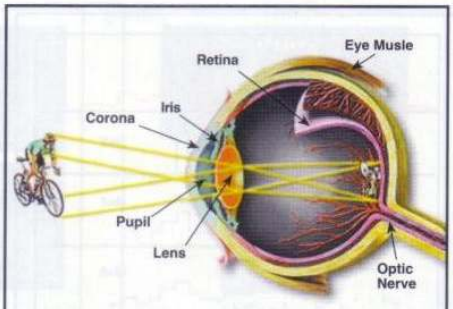
أهمية الرؤية الروبوتية

يكثر استخدام منظومات الرؤية الروبوتية في الروبوتات الصناعية مثل روبوتات التجميع والطلاء واللحام وتناول المواد والتغليف.. كما تستخدم الرؤية الروبوتية في عملية الفحص لاكتشاف عيوب التصنيع أو التجميع لضبط الجودة. مثل اكتشاف عدم وجود أجزاء ضرورية في المنتج النهائي أو وجود أجزاء مكسورة أو في مكان غير صحيح أو وجود أخطاء في شهادة مواصفات المنتج المطبوعة بطريقة أتوماتيكية كما هو الحال في مصانع الأدوية. كما تساعد الرؤية الروبوتية على القيام بأعمال الصيانة في الأماكن التي يصعب أو يستحيل على الإنسان التواجد فيها.. مثل فحص العازلات الكهربية أو استبدال وصيانة كابلات الضغط العالي أو صيانة المحطات النووية.. إلخ. كما تستخدم الرؤية الروبوتية بشكل أساسي.. في روبوتات استكشاف الفضاء والروبوتات البحرية وروبوتات البحث والإنقاذ التي تعمل في حالات الكوارث للبحث عن وجود أحياء في الأماكن التي يصعب الوصول إليها. ولعلنا نذكر المحاولة التي قامت بها National Geo-graphic لاستكشاف هرم خوفو (العدد ٧٤) بالاعتماد على منظومة رؤية روبوتية. وتستخدم هذه المنظومات أيضًا وبكثرة.. في مصانع الالكترونيات في عمليات تجميع وفحص الدوائر المتكاملة.. وفي التطبيقات الزراعية – في بعض روبوتات جني الفاكهة – لتحديد مكان الثمرة وتصنيف المنتجات الزراعية حسب اللون أو الحجم.

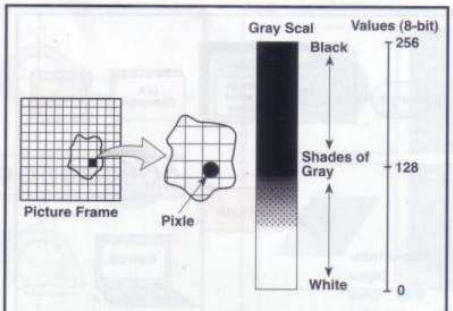
الصور الرقمية Digital Images

تعتمد جودة الصورة على قدرة الكاميرا على التمييز بين نقطتين متجاورتين تفصلهما مسافة بالغة الضآلة.. وهو ما يعرف باسم الوضوح Resolution. يتم تمثيل الصور الرقمية بواسطة مصفوفات ثنائية البعد 2D تحتوي على قيم تمثل شدة الضوء المنعكس Intensity.. وتحدد أبعاد المصفوفة الوضوح الفراغي للصورة Spatial Resloution.. بينما يحدد عدد الشفرات الرقمية المستخدمة في عناصر المصفوفة Data Bits وضوح الشدة In-tensity Resolution. تمثل كل خلية في المصفوفة عنصرًا من عناصر الصورة يسمى (Pixel: Picture Element). ويتم تمثيل الصورة الملونة بواسطة ثلاث مصفوفات من الخلايا تمثل إحداها شدة اللون الأحمر والثانية شدة اللون الأخضر والثالثة شدة اللون الأزرق (.RGB Sys). يمكن استخدام معاملات أخرى كالتشبع Saturation أو الضبابية Blur بدلاً من الشدة لتشكيل الصورة.

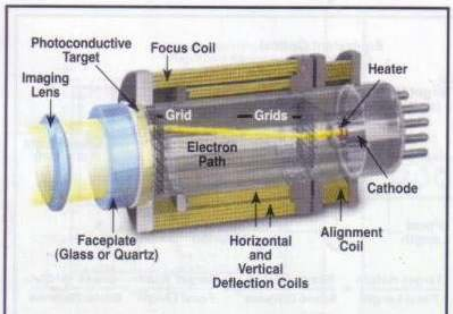
تحتوي معظم الصور الملونة ثنائية الأبعاد على مصفوفة تتكون من ٥١٢×٥١٢ عنصرًا.. يحتوي كل منها على قيمة بطول ٨ بت تمثل قيم الألوان الأحمر والأخضر والأزرق على الترتيب.. وبالتالي يكون حجم المعلومات التي



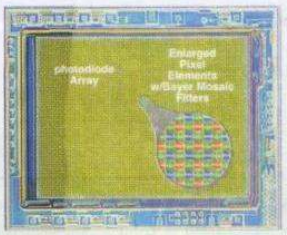
شكل رقم (١) : العين البشرية



شكل رقم (٢) : التدريج الرمادى



شكل رقم (٣) : كاميرا " Vidicon "

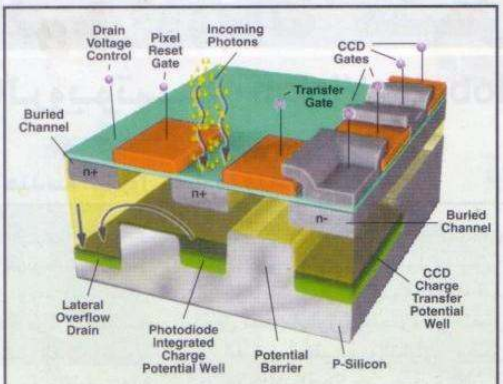


شكل رقم (٤) : مصفوفة CCD

تحتويها الصورة هو ٣×٥١٢×٥١٢× (٨بت). وفي حالة الصورة ثلاثية الأبعاد.. يتم تمثيل الصورة بواسطة مصفوفة حجمية ثلاثية الأبعاد 3D Voulmetric Arrays. وفي هذه الحالة.. تحتوي كل خلية من خلايا المصفوفة على عنصر حجمي يسمى فوكسل (Voxel: Voulme Element). وتصل السعة الإجمالية المطلوبة لتخزين الصورة إلى ٢ .. أي حوالي ٨ بليون بت في حالة تمثيل الصورة بمصفوفة ٥١٢×٥١٢×٥١٢.. واستخدام عناصر ٨ بت في كل خلية.. مما يجعل استخدام الصورة ثلاثية الأبعاد غير عملي بالرغم مما يوفره من معلومات تساعد على تحديد ماهية الأغراض الموجودة في الصورة.

وفي حالة استخدام بيانات مدى تمثل المسافة الفاصلة بين نقطة المشاهدة View Point والغرض (أنظر العددين ٧٨ و٨٠).. يتم تشكيل مصفوفة مستطيلة Rectangular Array من الخلايا يسمى كل عنصر فيها «رنجل» (Rangel: Range Element). وفي هذه الصورة.. تختفي قيم المدى لبعض الأغراض غير المرئية في المشهد والتي لا يمكن رؤيتها من نقطة مشاهدة معينة. لذلك.. لا يعتبر هذا التمثيل ثنائي الأبعاد 2D.. وفي نفس الوقت لا يمكن اعتباره أيضا ثلاثي الأبعاد 3D بشكل كامل لاختفاء بعض القيم.. لذلك يسمى تمثيل 2 1/2 D.

وبالإضافة إلى الوضوح.. تستخدم معاملات أخرى لتحديد جودة الصورة.. مثل خاصية التباين Contrast والتي



شكل رقم (٥) : الخلايا الحساسة للضوء " Photosites "

تعرف على أنها قدرة الكاميرا على استشعار الفرق في التظليل Shading بين بيكسل وآخر. وتستخدم أنواع كثيرة من الكاميرات الآن.. التدريج اللوني الرمادي.. حيث يتم استخدام من ٤-٨ درجات رمادية تتراوح ما بين الأبيض والأسود التام لكل بيكسل.. كما هو مبين بالشكل رقم (٢).

الكاميرات

توجد أنواع كثيرة من الكاميرات بنظريات عمل مختلفة يمكن استخدامها في التطبيقات الروبوتية.. نعرض منها النوعين التاليين:

– كاميرات Vidicon:

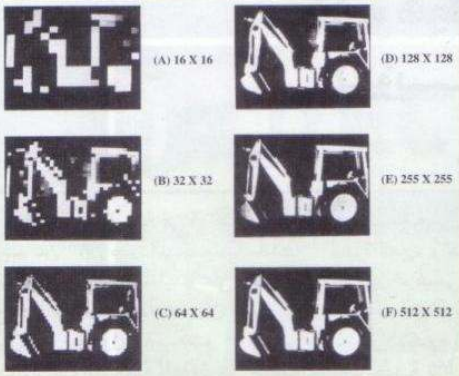
استخدمت هذه الكاميرات في السابق في كثير من التطبيقات الروبوتية نظرًا لانخفاض تكلفتها ودقتها المقبولة نوعًا ما – شكل رقم (٣). تتشابه فكرة عمل هذه الكاميرات مع فكرة عمل جهاز التليفزيون.. حيث تحتوي على انبوبة مفرغة من الهواء وكاثود وشبكة تحكم بالإضافة إلى شبكة من المكثفات الضوئية الدقيقة موضوعة على لوح يعمل كآنود. وتفرغ هذه المكثفات شحنتها بمعدل يعتمد على الضوء الساقط عليها.

تنبعث حزمة من الالكترونات من الكاثود يتم تعجيلها بواسطة المجال الكهربي بجهد عال.. ويتم استخدام مجال مغناطيسي للتحكم في مسار هذه الحزمة. عند سقوط حزمة الالكترونات على شبكة المكثفات.. يتم إعادة شحنها حيث يتناسب تيار إعادة الشحن مع كمية الشحنة المفقودة والتي تتناسب مع شدة الضوء الساقط وبالتالي يمثل هذا التيار إشارة الخرج للكاميرا.

– كاميرات CCD:

تكتنف الكاميرات المعتمدة على استخدام الأنابيب المفرغة عيوبًا كثيرة.. من أهمها التشويه واستهلاك القدرة والحاجة إلى استخدام جهد عال بالإضافة إلى الحجم مما جعل كاميرات CCD تحل محلها في كثير من التطبيقات الروبوتية.

تحتوي الكاميرا CCD على قطعة صغيرة مستطيلة الشكل من السيليكون تسمى Charge Coupled Device (CCD) تم إنتاجها في السبعينيات في معامل شركة «بل» لاستخدامها كوحدة ذاكرة في الحاسبات.. وتستخدم الآن بكثرة في العديد من الأجهزة مثل الكاميرات والتلسكوبات وأجهزة الفاكس والماسحات الضوئية. تحتوي هذه القطعة كما هو مبين بالشكل رقم (٤) على مصفوفة من الخلايا الحساسة للضوء تسمى Photosites تعمل على تحويل الضوء المنعكس من الغرض المراد تصويره إلى شحنة الكترونية يتم تخزينها فيما يسمى ببئر الجهد Potential Well.. وبعد ذلك يتم نقلها خلال الشريحة باستخدام مسجلات Registers وتكبيرها باستخدام وحدات تكبير Amplifiers كما هو مبين بالشكل رقم (٥).

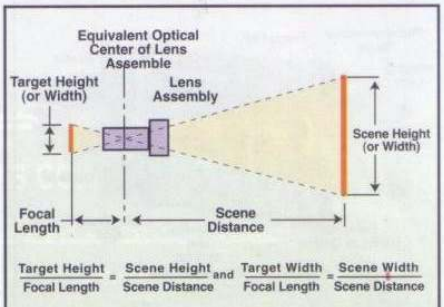


شكل رقم (٦) : تأثير عدد " البيكسلات "علي وضوح الصورة

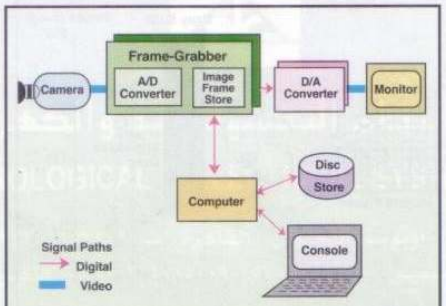
ويقاس وضوح الكاميرا Resolution بعدد العناصر (البيكسلات) المستخدمة في تشكيل الصورة. يوضح الشكل رقم (٦) تأثير عدد «البيكسلات» على وضوح الصورة.. حيث يلاحظ زيادة الوضوح بزيادة عدد «البيكسلات» المستخدمة.

ويجب الأخذ في الاعتبار هنا.. أن معايير الجودة للصورة التي يراها الروبوت تختلف اختلافًا كبيرًا عن تلك التي تقوم عليها جودة الصورة. على شبكية العين البشرية. وفي هذا الاتجاه.. نجد أن من أهم ما يميز منظومات الرؤية الروبوتية عن منظومة الرؤية البشرية. هو أن الروبوت يرى فقط ما ينفعه لأداء المهمة المكلف بها. فروبوت التجميع على سبيل المثال.. يجب عليه تحديد أركان القطعة المراد تجميعها دون الحاجة لمعرفة كل صفاتها. لذا.. نجد أنه لا يتم – بالضرورة – ترجمة الصورة المتكونة في منظومة الرؤية الروبوتية كما هي في العقل الروبوتي.. مثلما يحدث في منظومة الرؤية البشرية.. إذ أن الروبوت يستخدم فقط ما يفيد وحدة التحكم من مفردات الصورة كما سيأتي لاحقا.

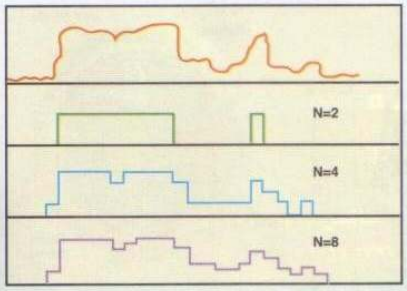
وبجانب الوضوح.. هناك الكثير من المعاملات التي تحدد مواصفات الكاميرا.. ومن أهمها الرقم البؤري f-number والذي يشير إلى سرعة العدسة وقدرتها على تجميع الضوء والذي يمكنه من تعيين قيمته من العلاقة f-number= Focal Length/Lens Diamter.. كما هو مبين بالشكل رقم (٧).



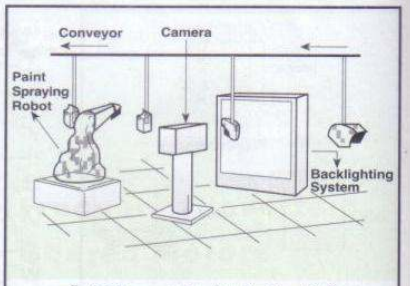
شكل رقم (٧) : الرقم البؤري



شكل رقم (٨) : اكتساب الصورة

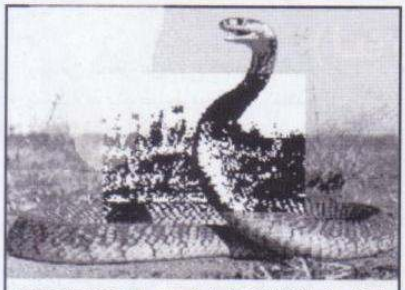


شكل رقم (٩) : الترقيم



شكل رقم (١٠) : الصور الظلية

شكل رقم (١١) : اكتشاف علامات معروفة في الصور


شكل رقم (١٢) : تحليل الصور بـ Thresholding

نظرية الإبصار الآلي
Machine Vision

في محاولة لمحاكاة منظومة الرؤية البشرية المعقدة.. تم تصميم كثير من منظومات الرؤية الروبوتية التي تعتمد على القيام بعمليتين أساسيتين.. هما اكتساب الصورة وتحويلها إلى صورة رقمية Image Acquisition.. ثم محاولة تحليل الصورة Image Analysis وتفسيرها وفهمها طبقًا للهدف المطلوب من منظومة الرؤية. ففي بعض الأحيان.. يكون الهدف هو التعرف على أبعاد الغرض أو تحديد مكانه أو لونه. وأحيانًا أخرى تهدف المنظومة إلى التعرف على عيوب التصنيع.. كما هو الحال في روبوتات الفحص وضبط الجودة. وتجدر الإشارة هنا.. إلى وجوب اختيار منظومة الرؤية الروبوتية منذ البداية طبقًا للتطبيق المطلوب.. فالمبالغة في درجة الوضوح لها مساوئها أيضاً.. فبالإضافة إلى أهمية توفير قدرة حاسوبية أكبر لمعالجة الصورة.. نجد أن الوضوح قرين التكبير.. وهو الذي يؤدي إلى إظهار بعض التفاصيل بصورة مبالغ فيها مثل خشونة السطح بين الكترودين في حالة روبوتات اللحام. وقد يؤدي إظهار هذه التفاصيل غير الضرورية.. إلى إعاقة العملية بدلاً من تحسينها.

- اكتساب الصورة Image Acuision:

يوضح الشكل رقم (٨).. العناصر الأساسية لمنظومة اكتساب الصور. وبالمقارنة بالعين البشرية كجهاز لالتقاط الصور.. تستخدم الكاميرات في منظومات الرؤية بالكمبيوتر لتكوين صورة ضوئية تناظرية للغرض المتعامل معه. وهناك أنواع كثيرة من الكاميرات كما أشرنا سالفًا.. مثل Charge Couple Device (CCD), Photo Diode Array or Charge Injection Device (CID).

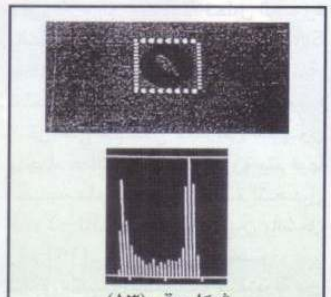
تتصل الكاميرا بالمنظومة الحاسوبية عن طريق Frame-grabber يحتوي على وحدة تحويل إشارات من تناظرية Analog إلى رقمية Digital.. بالإضافة إلى وحدة التخزين. وتتوقف القدرة الحاسوبية المطلوبة.. على الغرض من منظومة الرؤية وطبيعة الصور المطلوب معالجتها من حيث درجة الوضوح وما إذا كانت الصور ثنائية أو ثلاثية الأبعاد. تسمى عملية تحويل الصور التناظرية إلى رقمية.. بعملية الترقيم Digitization. يتم في هذه العملية تحويل الإشارة التناظرية التي تحتوي على عدد لانهائي من القيم إلى عدد صحيح من ١ إلى N.. حيث تمثل N درجة التدريج الرمادي الذي يمكن للمنظومة التعرف عليه. ويتم تخزين بيانات الصورة الرقمية في منظومة الرؤية باستخدام مسجلات تنائية Binary Registers. لذا.. نجد أن N تأخذ قيمًا أساسية للرقم ٢ مثلاً مثل ٢، ٤، ٨، ٦٤، ٢٥٦.. وهي التي تقابل مسجلات تنائية أطوالها ٢، ٣، ٦، ٨ على الترتيب. وهنا.. يجب ملاحظة أن المستوى الأول (N=2) لا يحتوي على إمكانيات رمادية.. ويحتفظ بمكانين فقط للونين الأبيض والأسود. وهناك العديد من التطبيقات الصناعية التي تستخدم هذا المستوى نظرًا لرخصه وبساطته وعدم احتياجه لأي سعة تخزينية كبيرة.. وهو ما يوضحه الشكل رقم (٩).. حيث يلاحظ أنه على الرغم من اختفاء عدد كبير من التفصيلات عند المستوى الأول (N=2) إلا أن هذا المستوى لا يشغل أكثر من ٢٥٪ من السعة التخزينية التي يتطلبها المستوى N=8.

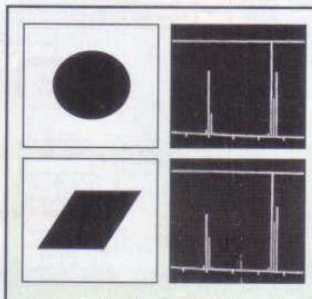
بالإضافة إلي ذلك.. تعتبر منظومة الإضاءة من الأمور المهمة الواجب دراستها بعناية في منظومات الرؤية. فلتكوين صورة جيدة.. يجب توفير إضاءة جيدة يختلف مفهوم جودتها عن المفهوم المعتاد بالنسبة للبشر. وفي معظم منظومات الرؤية الروبوتية.. لا يتم الاعتماد على الضوء الخارجي فحسب.. وإنما يستخدم مصدر إضاءة مستقل يمكن أن يكون مصدر ضوئي نقطي لتوضيح بعض التفاصيل في الغرض المتعامل معه أو حتى للاستفادة من الظلال الحادة في بيان هذه التفاصيل. وفي منظومات أخرى.. يتم إحاطة الغرض بعدد من المصادر الضوئية للتخلص تماماً من الظلال التي قد يؤدي تفسيرها إلي اكتشاف تفاصيل وهمية. وإذا كان المطلوب تحديد الإطار الخارجي للغرض أو حوافه دون الاهتمام بالتفاصيل الداخلية.. فيمكن استخدام طريقة الصور الظلية Silhouette.. وفيها يتم تسليط مصدر ضوئي من خلف الغرض للحصول على صورة سوداء مصمتة شديدة التحديد للغرض كما هو مبين بالشكل رقم (١٠). ولكن.. في حالة ضرورة الحصول على التفاصيل السطحية للغرض مثل المجاري والثقوب أو عيوب التصنيع كما هو الحال في روبوتات الفحص.. يتم استخدام إضاءة وجهية بدلاً من الإضاءة الخلفية.

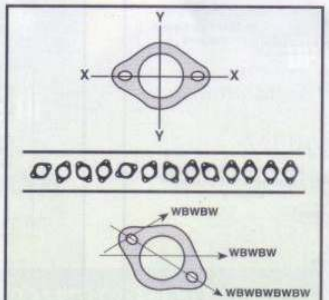
- تحليل الصورة Image Analysis:

لتحليل الصورة الرقمية بغرض تفسيرها وفهمها.. تجري الخطوات التالية:

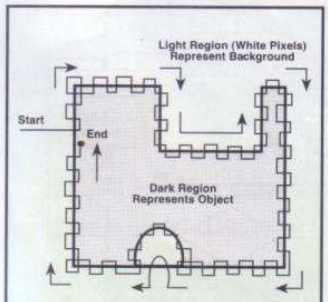
١- النوافذة Windowing: ويتم في هذه العملية التركيز على مساحة صغيرة أو نافذة في الصورة بغرض توفير وقت المعالجة والسعة التخزينية المطلوبة.. كما هو الحال في الإنسان عند تركيز بصره على مساحة محدودة في الصورة مع احتفاظه في الوقت نفسه ببقية الصورة في مجال رؤيته. وفي غالبية التطبيقات العملية .. تستخدم نافذة ثابتة Fixed Windows لا يتغير وضعها داخل الصورة. وفي هذه الحالة.. يجب مراعاة ضبط وضع الغرض المراد التعرف عليه داخل إطار النافذة. وفي منظومات الرؤية الأكثر تعقيدًا.. يتم استخدام النوافذ المكيفة Adaptive Windowin.. وفيها تكون المنظومة قادرة على اختيار النافذة المناسبة دون الحاجة إلى أية تجهيزات مساعدة لضبط وضع الغرض داخل إطار النافذة.. حتي يتم البحث في الصورة بأكملها بغرض اكتشاف علامات معروفة Landmarks تحدد وضع واتجاه الغرض. ويمكن عندئذ.. استخدام هذه العلامات لتحديد نطاق الاهتمام Area of Interest الذي سوف تشغله النافذة. يوضح الشكل رقم (١١) بعض العلامات التي تم استخدامها في منظومة الإبحار الطوبولوجي Topological Navigation لروبوت جوال.. حيث يلاحظ إمكانية استخدام كثير من العلامات مثل لوحة الاسم الموضوعة خارج المكاتب أو إشارات المرور أو علامات أجهزة الإطفاء أو مخارج الطوارئ إلخ.


شكل رقم (١٣)
طريقة الأعمدة البيانية "Histogramming"


شكل رقم (١٤)
شكل الجسم في التوزيع البياني



شكل رقم (١٥) : تحديد الشكل



شكل (١٦) : اكتشاف الحواف

وبغض النظر عن نوع النافذة (ثابتة أو مكيفة).. يتم ضبط مساحة النافذة حسب التطبيق.. حيث يمكن تصغيرها لتصبح «بيكسل» واحد إذا كان المطلوب معرفة لون أو درجة الرمادية لعينة منفردة.. إلا أنه يفضل من الناحية العملية استخدام عدة «بيكسلات» ثم حساب القيمة المتوسطة لدرجة الرمادية وذلك لتفادي فشل بعض القراءات بسبب عدام انتظام سطح الغرض أو بسبب مشاكل الإضاءة.

٢- البداءة Thresholding: بمجرد الانتهاء من تحديد نافذة الاهتمام في الصورة.. يمكن البدء في تحليل الصورة بهدف التعرف على الغرض الموجود في نطاق الاهتمام. تعتبر البداءة من أبسط أساليب تحليل الصور.. حيث يتم فيها اختزال الصورة إلى «بيكسلات» ثنائية بيضاء أو سوداء فقط للحصول على صورة ثنائية Binary Image. ولتوضيح هذه العملية.. يتم ضرب مثال طريف عن كيفية الإمساك بثعبان رمادي يزحف على أرضية من بلاط السيراميك الأبيض والأسود المصمم بنظام تبادلي مثل رقعة الشطرنج. في هذه الحالة.. يخصص «للبيكسلات» البيضاء والسوداء في الأرضية الرقم ١ (الأبيض).. أما «البيكسلات» التي تقع بين بداءتين Two Thresholds فتعطي الرقم صفر (الأسود) مما يجعل الصورة تتحول إلى ثعبان أسود على البلاط الأبيض مما يسهل الإمساك به. يوضح الشكل رقم (١٢) مثالاً مشابهًا للمثال السابق.. ولكن في هذه الحالة نجد أن الثعبان يتحرك في بيئة طبيعية مما يصعب عملية تمييزه. ويمثل الإطار الموجود في منتصف الصورة نتيجة عملية البداءة. نلاحظ.. أن هذه العملية تساعد على تمييز الغرض الموجود في نطاق الاهتمام عن الأغراض الأخرى أو الخلفية.

وفي التطبيقات الصناعية.. نجد أن الاختيار الجيد لنوعين من البداءة يساعد في الحصول على درجة رمادية مناسبة وبالتالي يمكن التقاط غرض محدد فقط مختلف في لونه عن الأغراض الأخرى أو عن الأرضية.

من الطرق الشائعة في اختيار البداءة.. طريقة الأعمدة البيانية Histro-gramming والتي تعتمد على إنشاء أعمدة تكرارية تمثل أعداد «البيكسلات» عند كل درجة رمادية داخل المنظومة. وإذا كانت الصورة تحتوي على غرض أسود على أرضية مضيئة.. نجد أن الأعمدة البيانية تكون ثنائية الهيئة Bi-modal – شكل رقم (١٣).

ويوضح الشكل رقم (١٤).. عدم اعتماد شكل التوزيع البياني التكراري على شكل الغرض. ومع ذلك.. يمكن استخدام هذه الطريقة في انتقاء الأشكال في حالة اختلاف درجة رمادية الشكل عن باقي الأشكال.. أو في حالة معرفة – ولو على وجه التقريب – عدد «بيكسلات» الشكل مقارنة بالأشكال الأخرى.

٣- تحديد الشكل Shape Identification: لمعالجة أوجه القصور في تحديد شكل الغرض باستخدام طريقة البداءة والأعمدة البيانية.. يتم استخدام خوارزمات مساعدة توفر معلومات أكثر عن شكل أو وضع الغرض الذي يتعامل معه الروبوت. فعلى سبيل المثال.. تم استخدام أحد خوارزمات التعرف على الأشكال لحل مشكلة تصحيح أوضاع مجموعة كبيرة من حابكات الزيت Oil Gaskets موضوعة على سير ناقل في خط إنتاج.. حيث يلزم توجيه القطر الأكبر لكل حابك على نحو يجعله موازيًا لمحور السير (y-y) – شكل رقم (١٥).

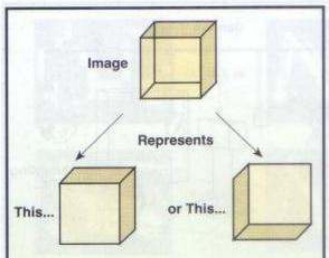
لإتمام هذه المهمة.. يقوم الخوارزم بعمل مسح عشوائي بواسطة مجموعة من الخطوط المستقيمة التي تخترق حدود الشكل الثنائي بزوايا مختلفة بحثًا عن أطوال المسارات Run-Lengths التي تصل بين «البيكسلات» البيضاء والسوداء في الشكل.. ليتم بعد ذلك تحديد المحور الأساسي للغرض (X-X) الذي يحتوي على التتابع التالي WBWBWBWBW.. حيث تشير W إلى اللون الأبيض و B إلى اللون الأسود.

٤- تحديد الحواف Edge Detection: يساعد تحديد الحواف (الكونتور) منظومة الرؤية على الحصول على معلومات أساسية عن ماهية ووضع الغرض المتعامل معه في حالة عدم الحاجة إلي معرفة التفاصيل الداخلية لهذا الغرض.. كما هو الحال في روبوتات التجميع وتداول المواد. يوضح الشكل رقم (١٦) مثالاً لعملية اكتشاف الحواف.. تظهر فيها صورة ثنائية يبدو بها الجسم معتمًا (ثنائي ١) والخلفية مضاءة (ثنائي صفر).

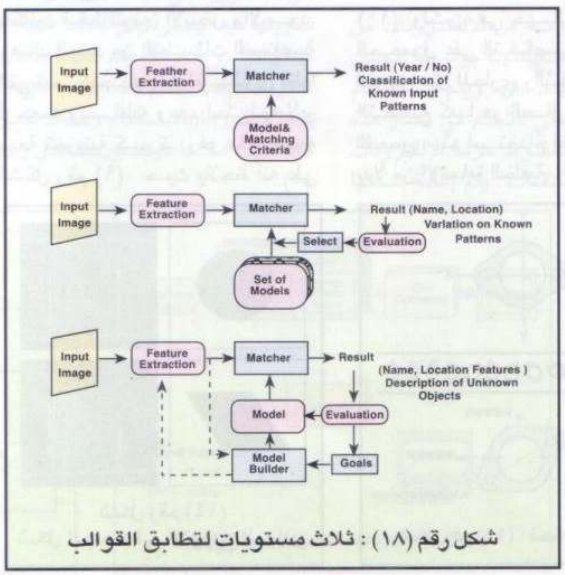
تبدأ العملية بمسار مستقيم عبر الخلفية لاختبار كل «بيكسل» على حدة لتحديد ما إذا كان معتمًا أو مضيئًا.. ويتصل المسار في حالة كون «البيكسل» مضيئًا. ويتم اعتبار أول «بيكسل» مضيء يتم اكتشافه كمرجع. وفي حالة اكتشاف «بيكسل» معتم – وهو ما يعني الدخول في المنطقة التي تمثل الغرض المراد اكتشاف حوافه – يتم تغيير مسار البحث والعودة مرة أخرى للبحث عن الحافة التي يدركها بمجرد اكتشاف «بيكسل» مضيئ. وتتكرر العملية حتى يتم الرجوع إلى «بيكسل» المرجع كما هو مبين في المخطط المنطقي – شكل رقم (١٧).

٥- تطابق القوالب Template Matching: تهدف هذه العملية إلى فهم الصورة للتعرف على الغرض المتعامل معه.. وذلك بمقارنة صورة الغرض المجهز بنموذج أو قالب معروف مسبقًا. يوضح الشكل رقم (١٨) ثلاثة من مستويات تطابق القوالب.. تختلف في قدرتها وفي طريقة معالجة الصورة حسب الغرض من عملية التطابق. ففي المستوى الأول.. نجد أن عملية المطابقة تهدف إلى استخراج الملامح الأساسية للصورة ومقارنتها بنموذج معروف لتحديد وجود تطابق من عدمه. يمكن استخدام هذا المستوى كعملية مبدئية في منظومة تجميع مثلاً لانتقاء الأجزاء الصحيحة المطلوب إدراجها إلى خط التجميع. في المستوى الثاني.. يتم استخدام مجموعة من النماذج – وليس نموذجًا واحدًا كما هو الحال في المستوى الأول – وعن طريق التغذية المرتدة بالإضافة إلى عدة قواعد للتقييم. يتم اختيار عدة نماذج في حالة عدم اكتشاف تطابق مقبول مع نموذج معين.. وتستمر العملية حتى يتم اكتشاف النموذج الذي يحقق أفضل تطابق. من الأمثلة العملية على ذلك.. استخدام الشبكات العصبية Neural Nets (العدد رقم ٧٥) في منظومات التعرف على الكتابة اليدوية.

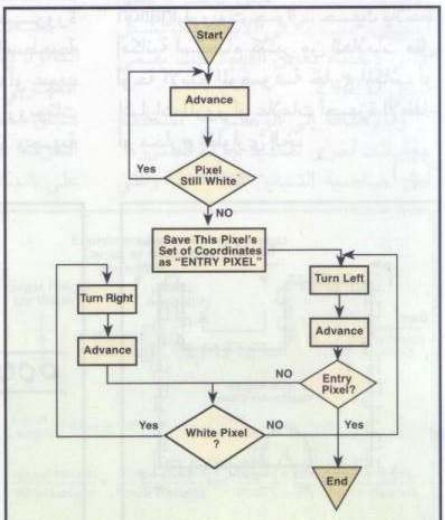
يعتبر المستوى الثالث أكثر المستويات تعقيدًا.. حيث يحتوي على عدد أكبر من مسارات التغذية المرتدة.. ويهدف إلى التعرف على صفات غرض مجهول تمامًا بدون وجود نموذج. لذا.. يتم استخدام وحدة إضافية تسمى بناء النماذج Model Builder.. وتتحول مهمة المنظومة من مجرد تحليل الصوة بالتطابق مع نموذج أو عدة نماذج معروفة – إلى محاولة تحليل الصورة بالتطابق مع نموذج يتم تخليقه Synthesis ذاتيًا بواسطة نفس المنظومة. تعتمد فكرة العمل.. على تخزين نموذج افتراضي مبدئي للغرض المجهول وإجراء عمليات مقارنة متكررة يتم فيها استخدام التغذية المرتدة لتعديل النموذج المفترض كما هو مبين بالشكل رقم (١٩).. حتى يتم الحصول على تطابق بنسبة سماحية مقبولة بين الغرض المجهول والنموذج الذي تم بناؤه بواسطة المنظومة.



شكل رقم (١٩) : النموذج الافتراضي



شكل رقم (١٨) : ثلاث مستويات لتطابق القوالب



شكل رقم (١٧) : المخطط المنطقي