تحسين شرح الصور التلقائي باستخدام قواعد الربط

ايمان صالح حمد الأحمدي

بحث مقدم لنيل درجة الماجستير في العلوم
(تقنية المعلومات)

إشراف
د. منيرة طيلب

كلية الحاسبات وتقنية المعلومات
جامعة الملك عبد العزيز
جدة – المملكة العربية السعودية
ربيع الثاني 1440هـ – ديسمبر 2018 م

تحسين شرح الصور التلقائي باستخدام قواعد الربط

ايمان صالح حمد الأحمدي

المستخلص

في الآونة الأخيرة، أصبحت أنظمة استرجاع الصور بصفة عامة وأنظمة استرجاع الصور بناء على المحتوى (CBIR) بصفة خاصة مجال بحث هام يُستخدم في مختلف المجالات. منذ الأيام الأولى، وأنظمة CBIR تعاني من وجود "مشكلة الفجوة الدلالية" الذي هو عدم وجود تطابق بين خصائص الصورة البصرية وبين النتائج التي يريدها ويتوقعها المستخدم. هذه المشكلة تؤثر بشكل رئيسي على انظمة تسمية الصور الأتوماتيكية. حيث أنه في الوقت الراهن استرجاع الصور بشكل سريع وفعَّال ليست مهمه سهلة, خصوصاً مع الزيادة المستمرة للصور الرقمية المخزنة على مواقع شبكات التواصل الاجتماعي.
أهم الدوافع لتنفيذ هذه الدراسة بالإضافة للحجم الهائل من الصور في الوسائط الرقمية هو أن دمج عدة وسائط يعتبر اتجاه جديد في أنظمة استرجاع الصور و تسمية الصور ويحتاج لمزيد من الدراسة. وحيث أن الصور في تطبيقات مختلفة تتواجد مع نصوص كتابية لها علاقة بالصورة فلا يوجد سبب لعدم استخدام كل المعلومات المتوفرة عن الصورة ودمجها لتنفيذ نظام تسمية يوفر المعنى الدلالي للصورة من خلال استخدام صفاتها المرئية والنصية معاً.
إذن فالمشكلة الأساسية هي كيف يتم التكهن بالمعنى الدلالي للصورة من خلال صفاتها المرئية والنصية. غالبية الطرق الحالية تفتقر للقدرة على استخراج المعنى الدلالي من الصورة بفاعلية. لذلك فهناك حاجة لدراسة طريقة الاستعلام المناسبة للمستخدم وكيفية استخدامها لاقتراح تسميات مناسبة وصحيحة للصور الرقمية اللتي لاحقاً سوف تسهل عملية استراجاعها.
الهدف الرئيسي من هذه الرسالة هو تصميم وتنفيذ نظام يساهم في سد الفجوة الدلالية في أنظمة تسمية الصور العامة كالصور المتواجدة في الشبكة العنكبوتية ومواقع التواصل الإجتماعي كـ Flickr. أيضا من اهم اهداف الرسالة دراسة الأعمال الحالية في مجال دمج المعلومات في تطبيقات الوسائط المتعددة بصفة عامة وفي أنظمة تسمية الصور بصفة خاصة ودراسة نقاط القوة والضعف في كل طريقة. نريد بالإضافة لذلك دراسة الطريقة المقترحة والتحقق من فاعليتها بمقارنتها مع مجموعات مشاركين استخدموا نفس قواعد البيانات في تجاربهم.
الأسلوب المقترح لدمج معلومات الصورة النصية والمرئية لأنظمة تسمية الصور - والتي هي الاتجاه الحديث في أبحاث تسميات الصور- يجمع بين اثنين من تقنيات تنقيب البيانات لاسترداد الصور ذات الصلة لغويا: خوارزمية منجم قواعد التجميع (ARM) وخوارزمية المجموعات (clustering) . منجم قواعد

التجميع الدلالي يتم إنشاؤه في المرحلة الأولى training حيث يتم في هذه المرحلة اكتشاف قواعد الارتباط بين مجموعات الصور بناء على العلاقات اللغوية للنص و مجموعات الصور بناء على المحتوى البصري. هذا المنجم يتم حفظه لاستخدامه لاحقا في مرحلة تسمية الصور annotation.
للتأكد من فاعلية النظام المقترح، تم تنفيذ النظام بإستخدام لغة C#.NET وباستخدام العديد من الأدوات لاستخراج الصفات المرئية والنصية للصور و لتقسيم الصور الى مجموعات clusters ثم أجريت التجربة على 8000 من صور ImageCLEF 2011 Flickr في المرحلة الأولى , وتم اختبار النظام على 10000 صورة في المرحلة الثانية. كما أنه تم اجراء التجربة على 15000 صورة من صور ImageCLEF 2012 في مرحلة التدريب ومن ثم اختبار النظام على 10000 صورة في مرحلة التسمية. وقد تم مقارنة نتائج مع نتائج مجموعة مشاركين استخدموا نقس قواعد بيانات الصور(training and testing dataset) في تجربتهم. كما أظهرت النتائج المتحصل عليها أن الطريقة المقترحة قد حققت أفضل قيمة للدقة Mean interpolated Average Precision and F-meauser بين فئات المشاركين.

Enhancement of Automatic Image Annotation Using Association Rules

Eman Saleh Hamad Alahmadi

ABSTRACT

Effective and fast retrieval of images from an image database is not an easy task, especially with the continuous and fast growth of digital images on the web. Images can be retrieved based on text describing or surrounding them, or based on the visual features. In both approaches, the retrieved images are not always semantically related to the query image. The automatic image annotation (AIA) is an approach that has been proposed and used by researchers to facilitate the retrieval of images semantically related to a query image. The AIA can be achieved with different approaches, text-based approach, content-based approach and the multimodal approach when the two modalities (text and visual features) are used to annotate the image.

A novel multimodal image annotation method is proposed in this thesis. The main purpose is to enhance the multimodal image annotation by taking advantages of textual and visual information and combine them using the association rules (ARs). The proposed method relies on clustering to regroup the text and visual features into clusters and on association rules mining to generate the rules that associate text clusters to visual clusters. The semantic ARs are generated during the training phase and later used in the annotation phase of the system. These rules semantically relate text clusters with visual clusters to predict a list of tags for each query image.

In the experimental evaluation, two datasets of the photo annotation tasks are considered; ImageCLEF 2011 and ImageCLEF 2012. The results achieved by the proposed method are better than all the multimodal methods of participants in ImageCLEF 2011 photo annotation task and state-of-the-art methods. Moreover, the MiAP of the proposed method is better than the MiAP of 7 participants out of 11 when using ImageCLEF 2012 in the evaluation.