Interest Pixel Mining

Qi Li
Western Kentucky University, USA

Jieping Ye
Arizona State University, USA

Chandra Kambhamettu
University of Delaware, USA

INTRODUCTION

Visual media data such as an image is the raw data representation for many important applications, such as image retrieval (Mikolajczyk & Schmid 2001), video classification (Lin & Hauptmann, 2002), facial expression recognition (Wang & Ahuja 2003), face recognition (Zhao, Chellappa, Phillips & Rosenfeld 2003), etc. Reducing the dimensionality of raw visual media data is highly desirable since high dimensionality may severely degrade the effectiveness and the efficiency of retrieval algorithms. To obtain low-dimensional representation of visual media data, we can start by selecting good low-level features, such as colors, textures, and interest pixels (Swain & Ballard 1991; Gevers & Smeulders 1998; Schmid, Mohr & Bauckhage 2000).

Pixels of an image may hold different interest strengths according to a specific filtering or convolution technique. The pixels of high interest strengths are expected to be more repeatable and stable than the pixels of low interest strengths across various imaging conditions, such as rotations, lighting conditions, and scaling. Interest pixel mining aims to detect a set of pixels that have the best repeatability across imaging conditions. (An algorithm for interest pixel mining is called a *detector*.) Interest pixel mining can be formulated into two steps: i) interest strength assignment via a specific filtering technique; and ii) candidate selection. The second step, candidate selection, plays an important role in preventing the output of interest pixels from being jammed in a small number of image regions in order to achieve best repeatability.

Based on interest pixels, various image representations can be derived. A straightforward scheme is to represent an image as a collection of local appearances—the intensities of neighboring pixels—of interest pixels

(Schmid & Mohr 1997). By ignoring the spatial relationship of interest pixels, this "unstructured" representation requires no image alignment, i.e., free from establishing pixel-to-pixel correspondence among imaging objects by image transformations such as rotation, translation, and scaling. Furthermore, the unstructured representation is very robust with respect to outlier regions in a retrieval application. However, the retrieval cost under unstructured representation is extremely expensive. In the context of face recognition, feature distribution is introduced to capture both global and local information of faces (Li, Ye & Kambhamettu 2006A). A limitation of feature distribution is the assumption of image alignment. A promising trend on interest pixel based representation is to build graph or tree representation for each image and measure the similarity of two images by the edit distance of their graphs or trees (Zhang & Shasha 1989). But as we will see in the later section, this trend is strongly supported by a recently proposed interest pixel mining method (Li, Ye & Kambhamettu 2008).

BACKGROUND

Most previous studies on interest pixel mining focus on exploring filtering techniques, i.e., the first step of an interest pixel detector, which leads to several widely-used filtering techniques such as gradient auto-correlation (Harris & Stephens 1988), and Difference of Gaussian (Lowe 2004). For the second step, previous studies usually adopt the so-called non-maximum suppression as the candidate selection scheme (Harris & Stephens 1988; Schmid, Mohr & Bauckhage 2000; Lowe 2004). Assume each image pixel has been assigned an interest strength. Non-maximum suppression

resets the strength of a pixel to zero, i.e., eliminates its candidacy, if it is not a local maximum. Non-maximum suppression is very effective in preventing interest pixels from being jammed in a small number of image regions. However, studies also show that non-maximum suppression over-suppresses good candidates of interest pixels if an image is weakly-textured, e.g., a face image. Besides the issue of over-suppression, non-maximum suppression may also destroy local geometry information (Li, Ye & Kambhamettu 2008). Brown et. al. proposed adaptive non-maximum suppression to obtain spatially well distributed interest pixels over images (Brown Szeliski & Winder 2005). Instead of using a fixed-size suppression window, adaptive non-maximum suppression dynamically decreases the size of the suppression window. Adaptive non-maximum suppression has been shown to be competitive to the standard one in image mosaicing.

MAIN FOCUS

Li, Ye & Kambhamettu (2008) proposed a novel candidate selection scheme, called imbalance oriented selection that chooses image pixels whose zero-/first-order intensities are clustered into two imbalanced classes (in size), for interest pixel mining. The basic motivation for imbalance oriented selection is to minimize the occurrences of edge pixels. (An edge pixel is a pixel on the boundary of an object or a scene.) It is worth noting that edge pixels are usually not good features in the context of image retrieval. This is because they have similar local appearances, which increases uncertainty in matching local appearances. Besides the issue of ambiguity, the existence of large numbers of edge pixels can also result in high computational cost in high-level applications. Figure 1 shows the key difference between non-maximum suppression and imbalance oriented selection, with a simple image: a white board with a black background. In this circumstance, the edge pixels on the four boundaries of the white board are expected to have distinctly larger strength than other pixels. Furthermore, due to the reality of the existence of noise, the strength of an edge pixel may be slightly different from the strength of its neighboring edge pixels. (The image of a white board in Figure 1 contains 0.1% noise.) So, after non-maximum suppression, the output of interest pixels consists of some edge pixels scattered around the four boundaries. However, with an imbalance oriented scheme, the output of interest pixels only consists of a few image pixels around the four corners of the white board, due to the stronger ability of the scheme in suppressing edge pixels.

FEATURE DISTRIBUTION

A feature distribution of an image represents the number of occurrences of interest pixels in regular grids of an image plane. Figure 2 shows the flow of generating a feature distribution for a simple input image. Feature distribution has been shown a good representation

Figure 1. 30 interest pixels in the image of a white board with a black background
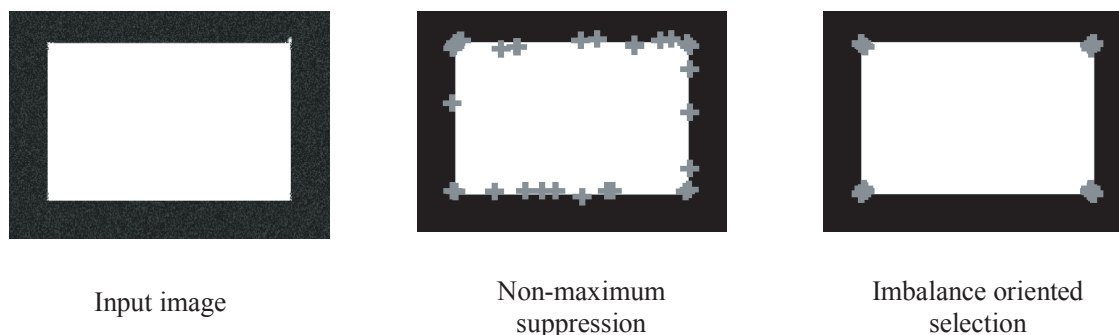
Input image

Non-maximum
suppression

Imbalance oriented
selection

Related Content

Place-Based Learning and Participatory Literacies: Building Multimodal Narratives for Change

Sharon Peckand Tracy A. Cretelle (2020). *Participatory Literacy Practices for P-12 Classrooms in the Digital Age (pp. 74-94).*

www.irma-international.org/chapter/place-based-learning-and-participatory-literacies/237415

Data Mining Applications in Steel Industry

Joaquín Ordieres-Meré, Manuel Castejón-Limasand Ana González-Marcos (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 400-405).*

www.irma-international.org/chapter/data-mining-applications-steel-industry/10851

Association Rules and Statistics

Martine Cadot, Jean-Baptiste Majand Tarek Ziadé (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 94-97).*

www.irma-international.org/chapter/association-rules-statistics/10804

Knowledge Discovery in Databases with Diversity of Data Types

QingXiang Wu, Martin McGinnity, Girijesh Prasadand David Bell (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 1117-1123).*

www.irma-international.org/chapter/knowledge-discovery-databases-diversity-data/10961

Web Design Based on User Browsing Patterns

Yinghui Yang (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 2074-2079).*

www.irma-international.org/chapter/web-design-based-user-browsing/11105