

Title	効果的な特徴マッチングのためのLocal Binary Pattern 特徴量に関する研究
Author(s)	Nguyen, Ngoc Thao
Citation	
Issue Date	2014-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/12295
Rights	
Description	Supervisor:宮田 一乗, 知識科学研究科, 博士

Abstract

“We are drowning in information but starved for knowledge.” (John Naisbitt, *Megatrends*, 1982) is a famous quote that best describes our status in this technology era. That is, the data is produced at an incredible rate while we have little ability to analyze it. The numerical and text data has been well treated in computers, thanks to their statistical analysis power granted by the data mining scientists. However, processing visual data has remained challenging since computers cannot interpret the data in the same manner as human. Therefore, how to represent visual data in computers and let them process it semantically has become a vital question nowadays.

The dissertation approaches the above question from a computer vision aspect. It searches for effective means to reconstruct the image properties in computers and hence improves the quality of the visual analysis. In this way, the information in images are correctly transformed from the non-declarative form to the declarative one, allowing us to manipulate the data easily on computers and gain insights into complex problems. This is expected to bring human toward the discovery of valuable visual knowledge.

The Local binary pattern (LBP) is studied comprehensively to enhance the representation of image properties. It is a popular family of visual features in computer vision due to their high discriminative power and computational simplicity. Specifically, the effectiveness of LBP features in three important tasks, including interest region description, pedestrian detection and background subtraction, is thoroughly examined. This demonstrates the robustness and generality of LBP in various visual tasks and thus well facilitating computers in interpreting different kinds of visual information.

The first contribution of this dissertation is the two novel LBP features for describing salient regions in images. Their underlying concepts are straightforward and computationally simple, thus they are suitable for many types of applications, especially those emphasizing the processing speed. The features are robust to different photometric and geometric image transformations, enabling us to achieve accurate correspondences between parts of images. These properties help the proposed interest region descriptors address two competing factors, matching accuracy and computational cost, at the same time, while this is still a critical issue for several modern descriptors using other features, such as gradient or pixel intensity. The success of novel LBP features motivates their development in higher-level tasks such as pedestrian detection, background subtraction and panorama image stitching.

The second contribution is a robust pedestrian detector using the proposed LBP feature in the first contribution. The encoding of this LBP is revised to better characterize edges along diagonal and vertical directions, which are most visible and meaningful details in an upright pedestrian body, so that the feature well distinguishes pedestrians from other objects in the image. The proposed detector combines LBP with color channels and gradient histogram to represent the subjects in different aspects, namely texture, color and gradient changes in magnitude and orientation. The advanced learning framework of [34] is adopted to resolve the computational bottleneck in constructing the feature pyramid. In this way, the proposed detector effectively identifies subjects of various poses in different challenging environments while achieving a speed of 15.2 frame per second (fps). This encourages its implementation in practical systems like surveillance, car driver assistant and human-computer interaction.

The third contribution is a multi-layer background modeling framework to extract moving objects from a video sequence. This framework models the scene background by processing consecutive frames through two layers: the block-wise layer considers blocks of pixels while the pixel-wise layer manipulates each individual pixel. The proposed LBP feature in the first contribution is used to represent the texture in a block, thus the framework is more robust to illumination changes and shadows, which frequently occur in background subtraction. The multi-layer framework operates in a coarse-to-fine manner to better reduce the errors than the traditional Mixture of Gaussians approach. It supports the object analysis for surveillance, video segmentation and event detection.

The final contribution is the introduction of an effective surveillance system that automates the detection of pedestrians in the monitored area. The three proposed techniques have been integrated perfectly to produce a unified system. The proposed system first extracts foreground regions using the background modeling framework in the third contribution then finds pedestrians in these regions with the pedestrian detector in the second contribution. When multiple cameras are used, a panorama view is created with the help of LBP features in the first contribution. This demonstrates the generality of the proposed LBP in the sense that they can participate effectively in all phases of the surveillance process. This doctoral research has been successful in improving the visual perception of computers to a semantic level. It therefore contributes to the computer vision as well as the knowledge discovery in Knowledge science.

Keywords

Feature matching, local binary pattern, pedestrian detection, background subtraction, surveillance