

| | |
|--------------|---|
| 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 |

| | | | |
|---------|---|--------------|------------------|
| 氏名 | NGUYEN NGOC THAO | | |
| 学位の種類 | 博士(知識科学) | | |
| 学位記番号 | 博知第 157 号 | | |
| 学位授与年月日 | 平成 26 年 9 月 24 日 | | |
| 論文題目 | A Study of Local Binary Pattern Features for Effective Feature Matching (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) | | |
| 論文審査委員 | 主査 | 宮田 一乗 | 北陸先端科学技術大学院大学 教授 |
| | | 吉田 武稔 | 同 教授 |
| | | 藤波 努 | 同 教授 |
| | | Dam Hieu Chi | 同 准教授 |
| | | 藤代 一成 | 慶応義塾大学 教授 |

論文の内容の要旨

“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 computationally 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

論文審査の結果の要旨

画像認識技術は、実世界とコンピュータの人工世界とをつなぐ技術として歴史が古く、いかにして効率的かつ効果的に画像の特徴量を抽出するかが重要とされている。本論文では、当該分野に関する綿密な文献調査に基づき、画像の特徴量に対する独自のアルゴリズムを提案し、その有効性および適用可能性を実証している。

論文では、画像の特徴量の抽出法として、画像内の局所的な特徴量(Local Binary Pattern)に注目し、画像の同定を行っている。これにより、画像から背景だけを抜き取り移動している物体を抽出する背景差分や、歩行者の抽出、また、複数画像を一枚の画像に統合するパノラマ画像の合成を効率的に処理することを可能にした。

提案手法は、効率的すなわち少ない計算量で的確に特徴量を捉えており、世界標準として公開されている画像群に対して、既存手法よりも堅牢かつ精度の高い結果を示している。また、提案手法は汎用性が高く、かつ計算量も少なく済むので、実用性が極めて高い。実験結果では、現状の標準的な監視カメラから得られる解像度(768×576)の画像に対して、標準的なスペックのデスクトップ PC 上で毎秒 15 フレームほどの処理速度で歩行者の抽出を可能としている。これは、監視カメラの記録レートに十分に見合ったパフォーマンスであり、リアルタイム処理が実現できている。提案しているアルゴリズムは並列処理が可能であり、GPU を用いたハードウェア上に実装する(GPGPU)ことで、更なる高いパフォーマンスを示すことも期待できる。

論文では、提案手法を組み合わせることで、複数の監視カメラから得られた画像から監視領域の広い画像を構築し、その画像から不審者を抽出したり、ショッピングモール内での買い物客の行動分析に利活用できる可能性を提示している。安全で安心な社会の構築や、消費者の行動分析によるマーケティング戦略などの基盤技術として、提案手法の多大な社会的貢献が十分に期待できる。国際会議での評価も高く、本研究成果に対して、某企業からオファーをいただいている。

以上、本論文は、画像内の Local Binary Pattern に注目して効率的かつ効果的な画像特徴量の抽出法を提案し、その有効性を膨大なデータに基づいて証明したものであり、学術的に貢献するところが大きい。よって博士(知識科学)の学位論文として十分価値あるも

のと認めた。