Pattern Recognition 42 (2009) 941 – 953

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

MSLD: A robust descriptor for line matching

Zhiheng Wang*,[1], Fuchao Wu, Zhanyi Hu

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 100080 Beijing, China

ARTICLE INFO

Article history:
Received 27 March 2008
Received in revised form 26 August 2008
Accepted 27 August 2008

Keywords:
Line matching
Line descriptor
MSLD descriptor

ABSTRACT

Line matching plays an important role in many applications, such as image registration, 3D reconstruction, object recognition and video understanding. However, compared with other features (such as point, region matching), it has made little progress in recent years.

In this paper, we investigate the problem of matching line segments automatically only from their neighborhood appearance, without resorting to any other constraints or priori knowledge. A novel line descriptor called mean–standard deviation line descriptor (MSLD) descriptor is proposed for this purpose, which is constructed by the following three steps: (1) For each pixel on the line segment, its pixel support region (PSR) is defined and then the PSR is divided into non-overlapped sub-regions. (2) Line gradient description matrix (GDM) is formed by characterizing each sub-region into a vector. (3) MSLD is built by computing the mean and standard deviation of GDM column vectors. Extensive experiments on real images show that MSLD descriptor is highly distinctive for line matching under rotation, illumination change, image blur, viewpoint change, noise, JPEG compression and partial occlusion.

In addition, the concept of MSLD descriptor can also be extended to creating curve descriptor (mean–standard deviation curve descriptor, MSCD), and promising MSCD-based results for both curve and region matching are also demonstrated in this work.

1. Introduction

Feature matching has drawn a lot of attention in the last few years. Great progress has been made and various approaches have been proposed for wide baseline point [1,2] and region [3] matching. Most of such approaches characterize local regions into feature descriptors and perhaps the most famous one is SIFT descriptor [1]. Line matching also plays an important role and is irreplaceable in many scenes [4–7]. A typical example is man-made scenes, which mainly consist of line segments, and line matching often becomes an unavoidable step for their 3D reconstruction. Unfortunately, compared to point and region matching, line matching is rarely reported in the literature and is still a challenging task due to various reasons [8]: inaccuracy of line endpoint locations, no strong disambiguating geometric constraint available, lacking of rich textures in line local neighborhood and so on.

Only a few methods are reported in the literature. In Ref. [9] the trifocal tensor is used for line matching by finding lines satisfying geometrical constraint in three views. Schmid and Zisserman [8,10] takes the epipolar constraint of line endpoints for short baseline matching, and one parameter family of plane homographies for wide baseline matching. However, both the trifocal tensor method [9] and the epipolar methods [8,10] demand known geometrical relations between images in advance. Lourakis et al. [11] present an approach using the "2 lines + 2 points" projective invariant for line matching in images of planar surfaces, and hence their method is limited to planar scenes. Herbert [12] proposes a method for automatic line matching in color images, where an initial set of line segment correspondences are generated using color histogram, then a topological filter is used to iteratively increase possible matches. The main drawback of this method is its heavy reliance on color rather than purely on texture. While color provides a very strong cue for discrimination, it may fail in the case where color feature is not distinctive, such as in gray images or remote sensing images. Although grouping matching strategy [13] has the advantage that more geometric information is available for removing ambiguities, and is able to cope with more significant camera motion, it often has high computational complexity and is sensitive to line topological connections or inaccuracy of endpoints.

As said in the above, most existing methods in the literature either require some prior knowledge [8–10] or are limited to some specific scenes [11,12] or are of high computational complexity [13], a method capable of automatically matching lines in general scene is

* Corresponding author.
E-mail address: zhwang@nlpr.ia.ac.cn (Z. Wang).

[1] Supported by the National Natural Science Foundation of China (60575019) and the National High-Tech Research and Development Program of China (2006AA01Z116 and 2007AA01Z341).

0031-3203/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.08.035

needed. This paper present a novel descriptor called mean–standard deviation line descriptor (MSLD) for automatic line matching without resorting to any prior knowledge. Compared with previous approaches, MSLD has two major appealing characteristics: one is that it is purely image content-based and can work without any other possible constraints, and the other is that MSLD is applicable to general scenes while some state-of-art methods are only scene-specific. Experiments show that MSLD descriptor is highly distinctive and robust against image rotation, illumination change, image blur, viewpoint change, noise, JPEG compression and partial occlusion.

The remainder of this paper is organized as follows. Section 2 introduces the definition of pixel support region (PSR) and the strategy of partitioning PSR into sub-regions, which is the key step to create line descriptor. In Section 3, each sub-region is characterized by a feature vector using a SIFT-like strategy. Section 4 elaborates the construction of MSLD descriptor. In Section 5, descriptor dimension and matching criteria are investigated, and Section 6 is the experiments. In Section 7, MSLD is extended for curve and region matching. Section 8 lists some concluding remarks.

2. Pixel support region (PSR)

Similar to creating point descriptor, how to select and partition line local neighborhood is the first step to construct our line descriptor. In this paper we propose a novel scheme to summarize local neighborhood of different-length lines into uniform description vectors.

Given a line segment L, as shown in Fig. 1, firstly two directions are introduced before defining the PSR: the average gradient direction $\mathbf{d}_\perp$ of pixels on the line and its anticlockwise orthogonal direction $\mathbf{d}_L$. For each pixel on the line L, a rectangular region centered at it and aligned with the directions $\mathbf{d}_\perp$, $\mathbf{d}_L$ is defined as the PSR. The PSRs of the pixels on the line along the direction $\mathbf{d}_L$ are denoted as $G_1, G_2, \ldots, G_N$ (assuming L consists of N pixels). In order to give a more distinctive description for the PSR, each PSR is divided into M non-overlapped sub-regions with the same size along the direction $\mathbf{d}_\perp$: $G_i = G_{i1} \cup G_{i2} \cup \cdots \cup G_{iM}$. It is noted that using $\mathbf{d}_\perp$ is necessary for the final descriptor to be rotation invariant, otherwise, some ambiguity may occur when deciding the order of $G_{i1}, G_{i2}, \ldots, G_{iM}$.

3. Sub-region representation

In this section, each sub-region will be characterized by a description vector using a SIFT-like strategy.

At first, it is noted that the gradient vector is not rotation invariant: suppose two images are related by $h(x') = f(\mathbf{R} \cdot x)$, the gradient vectors $\nabla \mathbf{f}$, $\nabla \mathbf{h}$ of a pair of corresponding points in the two images

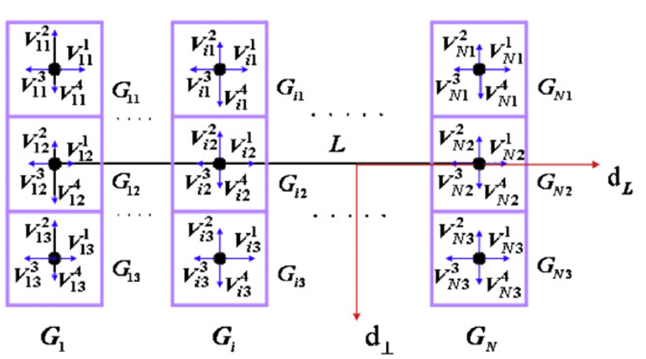

Fig. 1. Schematic figure of MSLD construction. In this figure, each PSR is divided into three sub-regions for illustration purpose, whereas nine sub-regions are adopted in our work.

must satisfy the relation $\nabla \mathbf{h} = \mathbf{R} \cdot \nabla \mathbf{f}$. To achieve rotation invariance, each sample gradient is rotated aligned with the directions $\mathbf{d}_\perp$ and $\mathbf{d}_L$: $\nabla_L \mathbf{f} = (\nabla \mathbf{f} \cdot \mathbf{d}_\perp, \nabla \mathbf{f} \cdot \mathbf{d}_L)^{\mathrm{T}} \triangleq (f_{d_\perp}, f_{d_L})^{\mathrm{T}}$. This step is similar to that of SIFT, where the gradient orientations are aligned relative to the key-point orientation.

Motivated by SIFT, a Gaussian weighting function with a scale σ, equal to half of the PSR side along the direction $\mathbf{d}_\perp$, is used to assign a weight to each sample in the PSR: for a sample which has a distance d from the line L, its weight can be expressed as $w = 1/(\sqrt{2\pi}\sigma)\, e^{-d^2/2\sigma^2}$. The purpose of such a weighting is to give less importance on the gradients of those samples that are far from the line, which are most likely affected by mis-registration errors. Another reason is to reduce the descriptor's sensitivity to small change in the position of each PSR.

As one sample shifts smoothly from being within one sub-region to another, the descriptor may change abruptly and thus boundary effect arises. In order to reduce this effect, for a sample whose gradient is $\nabla \mathbf{f}$, it will contribute not only to its sub-region G_{ij}, but also to its nearest neighbor sub-region $G_{i(j-1)}$ (or $G_{i(j+1)}$) along the direction $\mathbf{d}_\perp$.

Denote the distances from it to the central lines (parallel to $\mathbf{d}_L$) of the two sub-regions are d_1, d_2, then the contributions to the two sub-regions are $\nabla \mathbf{f} \cdot w_1, \nabla \mathbf{f} \cdot w_2$, respectively, where $w_1 = d_2/(d_1 + d_2)$, $w_2 = d_1/(d_1 + d_2)$. This step of linear interpolation is only adopted along the direction $\mathbf{d}_\perp$ but not along the direction $\mathbf{d}_L$, because sub-regions are overlapped each other along the direction $\mathbf{d}_L$ and thus boundary effect is negligible.

Denote the gradients distributed in a sub-region G_{ij} as $\{(\tilde{f}_{d_\perp}, \tilde{f}_{d_L})^{\mathrm{T}}\}$, then a 4D feature vector is formed by accumulating these gradients along the directions $\mathbf{d}_\perp$, $\mathbf{d}_L$ and their opposite directions, respectively (as shown in Fig. 1):

$$\mathbf{V}_{ij} = (V_{ij}^1, V_{ij}^2, V_{ij}^3, V_{ij}^4)^{\mathrm{T}} \in R^4 \tag{1}$$

where

$$V_{ij}^1 = \sum_{\tilde{f}_{d_\perp} > 0} \tilde{f}_{d_\perp}, \quad V_{ij}^2 = \sum_{\tilde{f}_{d_\perp} < 0} -\tilde{f}_{d_\perp}, \quad V_{ij}^3 = \sum_{\tilde{f}_{d_L} > 0} \tilde{f}_{d_L}, \quad V_{ij}^4 = \sum_{\tilde{f}_{d_L} < 0} -\tilde{f}_{d_L}$$

It can be proved that $\mathbf{V}_{ij}$ constructed in such way is invariant to image rotation, and it is used as the description vector of the sub-region G_{ij}.

4. MSLD descriptor

By stacking the description vectors of all the sub-regions associated with a line segment, a $4M \times N$ matrix called line gradient description matrix (GDM) is formed as

$$\mathbf{GDM}(L) = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} & \cdots & \mathbf{V}_{1N} \\ \mathbf{V}_{21} & \mathbf{V}_{22} & \cdots & \mathbf{V}_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{V}_{M1} & \mathbf{V}_{M2} & \cdots & \mathbf{V}_{MN} \end{pmatrix} \triangleq (\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_N) \quad (\mathbf{V}_i \in R^{4M}) \tag{2}$$

GDM contains the most structural information in the line neighborhood region. However, it cannot be directly used as a line descriptor because its size still varies with line length. To make the descriptor independent of the line length, statistical entities from GDM column vectors are explored here. We have tested several popularly used statistic measures: mean, standard deviation, Fourier coefficients and moments. Among all the candidates and their combinations, we have found that the combination of mean and standard deviation can provide satisfying matching result, though the first four Fourier coefficients or moments can give slightly better performance. Considering descriptor's dimensional problem, the mean and standard deviation are adopted to construct our line descriptor