SHADOW EDGE DETECTION USING GEOMETRIC AND PHOTOMETRIC FEATURES

Arjan Gijsenij and Theo Gevers

Intelligent Systems Lab Amsterdam
Faculty of Science, University of Amsterdam
Science Park 107, 1098 XG Amsterdam, The Netherlands
{a.gijsenij, th.gevers}@uva.nl

ABSTRACT

The detection of shadow and shading edges is a first step towards reducing the imaging effects that are caused by interactions of the light source with surfaces that are in the scene. As most of the algorithms for shadow edge detection use photometric information, geometric information have been ignored so far. In this paper, the aim is to include geometric features for more robust shadow edge detection. First, thousands of patches are annotated as either containing a shadow edge or not. Then, geometric features of these patches are analyzed and it is shown that the combination of photometric and geometric features improves the classification of shadow edges with respect to using either one of these features with 14%. These results demonstrate the added value of geometric features, in addition to photometric features, for the detection of shadow edges.

Index Terms— Shadow edge detection, photometric features, geometric features

1. INTRODUCTION

Many image processing applications have to deal with imaging effects that are caused by interactions of the light source with surfaces that are in the scene. Especially for systems that have to work robustly in real-world complex scenes, there is a need to correct for effects like shadow and shading. A first step for reducing such effects is the detection of shadow edges. For instance, shadow removal can greatly benefit from accurate detection of shadow edges [1], and more accurate color constancy can be obtained using shadow edges than using material edges [2]. In this paper, the focus is on the classification of shadow and shading edges.

Shadows are regions in a scene which are not reached by the light source directly, e.g. because of occluding objects or the direction of the light source. Shadow regions have a lower and therefore contrasting intensity (and sometimes color) with respect to regions that are fully illuminated by the light source, and it is this property that is mainly exploited for the detection of shadow edges, e.g. [3, 4, 5, 6, 7]. For instance, an image that is invariant to the light intensity and color can be constructed by projecting the $2D$ log-chromaticity image onto the direction orthogonal to the light source [3]. Alternatively, images that are invariant to either specular edges, shadow edges, or specular *and* shadow edges can be obtained by subtracting the appropriate photometric variant from the derivative of an image [7].

As most of these shadow edge detection algorithms are based on photometric information, geometric features like SIFT [8] or local binary patterns [9] have been ignored so far. This is partly caused by the lack of ground truth: data sets that contain clear shadow edges are either small (e.g. the toy images shown in [7] or the limited number of images used in [1]) or focussed on photometric invariants (e.g. the collection of shadow and non-shadow patches introduced in [6]). In this paper, the aim is to include geometric information in the classification of shadow edges.

The contribution of this paper is to combine geometric and photometric information to detect shadow edges. First, thousands of patches are annotated and labelled as either containing a shadow edge or not. Then, the geometric features of these patches are analyzed and a classifier is trained to distinguish between shadow and non-shadow patches. The combination of photometric and geometric features will be exploited for classification of shadow edges in addition to using either photometric or geometric features alone. Experiments are conducted on a large set of manually annotated patches.

The rest of this paper is organized as follows. In section 2, the features that are used for classification of patches into shadow or non-shadow classes are discussed. Then, in section 3, the effectiveness of these features (and combinations of these features) is evaluated on a large set of manually annotated patches, followed by the conclusion in section 4.

2. SHADOW EDGE DETECTION

In general, shadows and shading regions are identified using photometric information. First, characteristics of shadows and shading are analyzed, resulting in classification scheme to identify different edges. However, analysis of shadow and

shading edges is merely done using low-level photometric properties, completely ignoring possible geometric regularities. In this section, several photometric invariants are shortly explained. After that, geometric features are introduced for the purpose of shadow edge detection.

2.1. Photometric features

Several approaches using photometric or color invariants have been proposed to detect shadow/shading edges, of which four types are explored in this paper. The main principle is to identify regions with a lower, and therefore contrasting, intensity with respect to regions that are fully illuminated by the light source.

Quasi-invariants. The quasi-invariants introduced in [7] are shown to be invariant to either shadow/shading edges, specular edges or both. These quasi-invariants can be obtained by subtracting the appropriate photometric variant from the derivative of an image. The photometric variants are called shadow-shading variant, specular variant and shadow-shading-specular variant, and by combing the three invariant images using a simple classification scheme, edges can be classified into material edges, shadow edges and specular edges.

Physics-based invariants. Invariants that are similar to the quasi-invariants are proposed in [4], where the invariants are derived from a physical reflectance model based on the Kubelka-Munk theory for colorant layers. Several invariants are proposed, but in this paper the combination of the E and C-invariant proved to perform best (although difference between the several invariants are small). The E-invariant is consistent with a grey-scale image, and therefore contains no invariance. The C-invariant is insensitive to the illumination direction and intensity. As features, the mean and standard deviation of the responses in the gradient of these invariants are computed for an entire patch.

Color constancy at a pixel. In [3], the focus is on obtaining an invariant image that does not depend on the light source (intensity nor chromaticity). By projecting the $2D$ log-chromaticity image onto the direction orthogonal to the light source [3], an invariant image is obtained that is insensitive to changes in intensity *and* color of the light source. The invariant image can be obtained as follows:

$$I = \cos(\theta) \cdot \log(\frac{R}{G}) + \sin(\theta) \cdot \log(\frac{B}{G}), \qquad (1)$$

where θ is based on the color of the light source. For the best result, the chromaticities of the light source should fall near the Planckian locus. Since the light source is generally unknown, the value for θ is fixed in this paper.

Normalized-rgb. The final color invariant that is examined in this paper is based on the normalized-rgb color space. This model is often used to compute color features, because of its invariance to intensity changes. It can be computed as follows:

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B}. \quad (2)$$

This color space is insensitive to changes in intensity, and should therefore be able to distinguish material changes (i.e. changes in color) from changes in intensity (i.e. shadow/shading edges). Consequently, shadow edges can be identified by taking the intersection of the edge-map of RGB-patches and the edge-map of rgb-patches. The edge-map of a single color channel is obtained using the Canny edge detector [5].

2.2. Geometric features

Three geometric features are selected, each describing a local region. The main idea is to analyze possible patterns in geometry that are characteristic for shadow edges. The three features are explained in short.

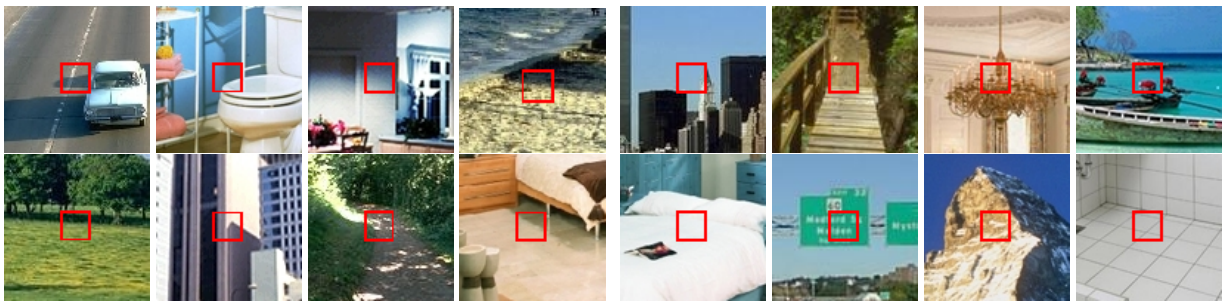
SIFT. The SIFT-feature [8] has proven to be very effective for object and scene recognition, and it describes the orientation of edge responses in a small region around a point. The original SIFT-feature is composed of a detector and descriptor phase. However, considering the size of the patches, only the descriptor is used to describe the content of a patch.

Local Binary Pattern. Local binary patterns [9] describe a texture using a histogram of binary patterns, where every binary pattern corresponds to one pixel in a region. The binary pattern for a pixel P describes the relative values of neighboring pixels: neighbors with lower values than pixel P are assigned the value 0, while the other neighbors (i.e. with higher or equal values) are assigned the value 1. By concatenating the values of all neighbors, a binary pattern for pixel P is obtained and the binary patterns of all pixels in a region are summarized into a histogram. The histogram of an entire patch is used as feature.

Grey-level Co-occurrence Matrix. The Grey-level co-occurrence matrix [10] is a well-studied texture feature, that captures the relationship between intensity values that occur simultaneously. The result is a matrix C, where element (i, j) indicates the number of times elements i and j co-occur in an image I at a given offset $(\Delta x, \Delta y)$. In this paper, 8×8 grey-level co-occurrence matrices are created by considering all 8-connected neighbors of a center pixel.

2.3. Combination

Combining features can be done using several approaches. In general, a distinction can be made between early fusion and late fusion. When early fusion is used, the feature vectors are merged (e.g. by concatenation) before learning a classifier. This often results in a feature vector with larger dimensionality. Late fusion results in a separate classifier for every feature, followed by a rule that combines the posterior probabilities, like the sum or product-rule. Experimental results show



(a) Patches with shadow edges

(b) Patches without shadow edges

Fig. 1. Examples of the shadow and non-shadow classes. Note that the actual patches that are used for the classification are marked by the red square, the rest of the image is shown to denote the context of the patch.

Feature	AUC	Error
Quasi-invariants	0.69	0.38
Physics-based invariants	0.72	0.42
Color constancy at a pixel	0.66	0.35
Normalized-rgb	0.70	0.39
SIFT	0.70	0.36
LBP	0.68	0.38
GLCM	0.70	0.36
All texture features combined	0.78	0.30
All photometric features combined	0.78	0.30
All features combined	0.82	0.26

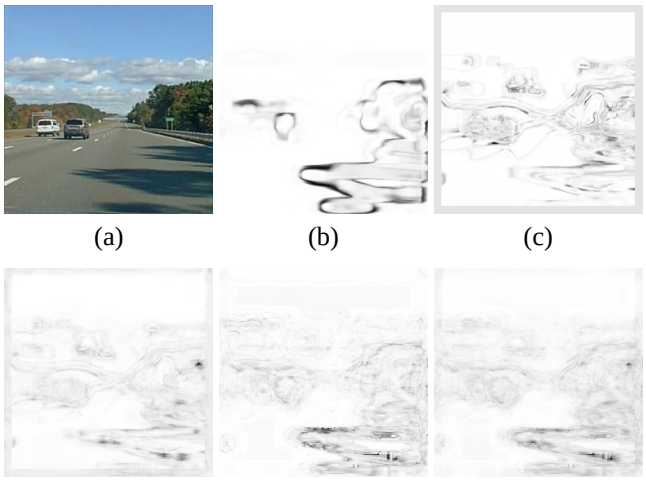
Table 1. Results of classification using several features. The AUC denotes the area under the corresponding ROC-curve, while the error denotes the ratio of misclassified patches when a test patch is classified using maximum posterior probability.

that late fusion combined with the sum-rule (i.e. the final posterior probability is the sum of the posterior probabilities of all separate classifiers) is the best choice for this problem [11]. Therefore, in this paper, late fusion will be adopted.

3. EXPERIMENTS

In this paper, shadow edge detection is performed on a data set containing manually annotated shadow and non-shadow patches. Shadow patches are patches with at least one clear shadow edge (and possibly other material edges). Non-shadow patches are patches with only material edges or no edges at all, see figure 1 for some examples. These patches are all 19×19 pixels and are extracted from both indoor and outdoor scenes. All original scenes are 256×256 pixels, and are taken from [12, 13]. In total, the set contains 7047 patches, of which 3461 are labelled as shadow patch, and 3586 are labelled as non-shadow patch. For the classification, a simple framework is used, i.e. 10-fold cross-validation using a 1-nearest-neighbor classifier. As performance measures, the area under the ROC-curve is computed, in addition to the classification error (i.e. the ratio of misclassifications using classification based on maximum posterior probability).

Photometric features. The performance of the four types of photometric features is reported in table 1. The area under the curve (AUC) is computed from the ROC-curves (not shown here), the classification error is computed from the number of misclassifications of all shadow and non-shadow patches. It can be seen that the physics-based invariants proposed in [4] performs best among the photometric features,



(a) (b) (c)

(d) (e) (f)

Fig. 2. An example of edge classification of the image in figure (a) the physics-based invariants or using co-occurrence matrices in figures (b) and (c). In figure (d), (e) and (f) results are shown when combination all texture features, all photometric features, or all texture *and* photometric features, respectively.

with the highest AUC and the lowest classification error. However, note that the results of [3] are obtained without calibration of the parameter θ. A fixed value is used, while the images are taken with multiple cameras. An example of using the physics-based invariants is shown in figure 2(b).

Geometric features. The performance of the geometric features are very similar to the performance of the photometric features. The best results when using geometric features are obtained using the SIFT and the grey-level co-occurrence matrices, see table 1 for results. An example using the grey-level co-occurrence matrices is shown in figure 2(c).

Combination. Using a combination of features can significantly improve the results. Fusing the output of all photometric features can boost performance to an area under the curve of 0.78 (an improvement of over 8%), and fusing the geometric features results in a similar performance. Using both photometric features *and* geometric features, however, performance is increased even further to 0.82, which is an improvement of 14% with respect to merely using photometric features (see table 1). Examples of combining different types of features are shown in figure 2(d)-(f).

3.1. Discussion

Even though the dimensionality of the photometric features is much lower than the dimensionality of the geometric features, the performance of the two features are very similar. Another advantage of the photometric features, is the possibility of an analytical approach to learning and evaluating an edge classifier. However, results in this section suggest that there is indeed information in the geometry of shadow and shading edges, that can benefit edge classification. Using both photometric *and* geometric features can significantly improve results, compared to merely using photometric or geometric features.

Analysis of the different geometric features shows that the most discriminating geometric features are diagonal edges or the complete lack of edges. The latter is explained by the fact that the class of non-shadow patches contains some patches without a clear gradient (e.g. uniformly colored regions like sky or road), while the class of shadow patches all contain, by definition, at least one clear gradient. The positioning of the light source can be an explanation for the discriminative power of diagonal edges. While an object itself might consist of horizontal, vertical and diagonal edges, the shadows that are caused by the light source are often diagonally oriented. For instance, a tall building with mainly horizontal and vertical edges (i.e. the shape of the building) will probably result in diagonal edges when looking at the cast shadow. Hence, a patch with a strong bias towards diagonal edges might be interpreted as a shadow patch based on these results.

4. CONCLUSION

In this paper, thousands of patches are annotated and labelled as either containing a shadow edge or not. After that, it is shown that geometric features have, to some extent, the ability to distinguish between shadow and non-shadow patches. The most discriminative features are diagonal edges and the lack of edges. Finally, it is shown that geometric features, in combination with more conventional features like photometric invariants can significantly improve shadow edge detection. The results presented in this paper clearly justify the need for further investigation into the use of geometric features for shadow edge detection.

5. REFERENCES

[1] G.D. Finlayson, S.D. Hordley, C. Lu, and M.S. Drew, "On the removal of shadows from images," *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 28, no. 1, pp. 59–68, 2006.

[2] A. Gijsenij, Th. Gevers, and J. van de Weijer, "Edge classification for color constancy," in *Proc. CGIV*, Terrassa, Spain, June 2008, pp. 231–234.

[3] G.D. Finlayson and S.D. Hordley, "Color constancy at a pixel," *J. Opt. Soc. Am. A*, vol. 18, no. 2, pp. 253–264, 2001.

[4] J.M. Geusebroek, R. van den Boomgaard A.W.M. Smeulders, and H. Geerts, "Color invariance," *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 23, no. 12, pp. 1338–1350, 2001.

[5] Th. Gevers and H. Stokman, "Classifying color edges in video into shadow-geometry, highlight, or material transitions," *IEEE Trans. Multimedia*, vol. 5, pp. 237–243, 2003.

[6] E.A. Khan and E. Reinhard, "Evaluation of color spaces for edge classification in outdoor scenes," in *Proc. ICIP*, 2005, pp. 952–955.

[7] J. van de Weijer, Th. Gevers, and J.M. Geusebroek, "Edge and corner detection by photometric quasi-invariants," *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 27, no. 4, pp. 625–630, 2005.

[8] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[9] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 24, no. 7, pp. 971–987, 2002.

[10] R.M. Haralick, K. Shanmugam, and I. Dinstein, "Textural geatures for image classification," *IEEE Trans. Sys. Man. Cyb.*, vol. 3, no. 6, pp. 610–621, 1973.

[11] J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Patt. Anal. Mach. Intel.*, vol. 20, no. 3, pp. 226–239, 1998.

[12] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *Int. J. Comput. Vision*, vol. 42, no. 3, pp. 145–175, 2001.

[13] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. CVPR*, 2006, pp. 2169–2178.