Constrained optical flow for aerial image change detection
Nicolas Bourdis, Marraud Denis, Hichem Sahbi

To cite this version:

HAL Id: hal-00667237
https://hal.archives-ouvertes.fr/hal-00667237
Submitted on 7 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ABSTRACT

Nowadays, the amount of video data acquired for observation or surveillance applications is overwhelming. Due to these huge volumes of video data, focusing the attention of operators on “areas of interest” requires change detection algorithms. In the particular task of aerial observation, camera motion and viewpoint differences introduce parallax effects, which may substantially affect the reliability and the efficiency of automatic change detection.

In this paper, we introduce a novel approach for change detection that considers the geometric aspects of camera sensors as well as the statistical properties of changes. Indeed, our method is based on optical flow matching, constrained by the epipolar geometry, and combined with a statistical change decision criterion. The good performance of our method is demonstrated through our new public Aerial Imagery Change Detection (AICD) dataset of labeled aerial images.

Index Terms— Change detection, aerial observation, constrained optical flow, epipolar geometry, parallax effects, AICD dataset.

1. INTRODUCTION

These last few years have witnessed a rapid development of video acquisition devices, which are becoming cheaper and more efficient. In the context of observation via airborne cameras, the volume of data being acquired is overwhelming human operators and most of it is stored unprocessed or lost. As a consequence, it becomes necessary to develop solutions addressing this bottleneck.

This paper addresses the specific problem of change detection in aerial images, and aims at focusing the operator’s attention on areas containing changes. Change detection [1] refers to the problem of detecting significant and possibly subtle changes between a reference and a new (called test) image (e.g. appearing or disappearing buildings or vehicles), while ignoring insignificant ones. General change detection is particularly challenging: difficulties arise when the reference and test images are acquired at distinct dates and with large viewpoint differences. Such conditions usually introduce many environmental changes (illumination, weather, ...)

and parallax effects due to scene content (trees, buildings, relief) and camera motion. Although a human expert may identify significant changes successfully, such conditions make the task extremely challenging to most of the current automatic approaches.

The method we present focuses on achieving resilience to viewpoint differences and parallax effects while remaining computationally efficient (involving no time-consuming search, for instance [2, 3]) and simple (involving no heavy scene model, like in [4, 5]) in order to allow for future real-time processing of aerial videos. For that purpose, we introduce a novel approach based on optical flow constrained with the epipolar geometry [6]. More precisely, our contribution is twofold: (i) the underlying mask of change is efficiently estimated using an optical flow algorithm constrained by epipolar geometry, and (ii) our change detection criterion is purely statistical and able to distinguish between real changes and parallax effects.

The remainder of this paper is organized as follows. First, section 2 explains why epipolar geometry provides an appropriate framework in order to cope with parallax effects. Then, section 3 emphasizes on our change detection algorithm, followed in Section 4 by the results obtained on an annotated data-set of aerial images. This benchmark will be publicly available as we believe such data-sets may play a great role in further developments and comparisons of efficient algorithms.
the Random Sample Consensus (RANSAC) algorithm. No-
scriptors and the epipolar geometry is then estimated using
both images, possible matches are identified using their de-
transformation. First, SURF keypoints are extracted from
image is registered to the reference one using a homography
aligned with respect to an arbitrary surface. A given test
As discussed in the previous section, images should be
should distinguish between real changes and parallax effects.
Many existing change detection criteria rely on simple image
differences [1], even though this makes true changes and par-
allax effects indistinguishable, as shown in Figure 1.

As demonstrated in [7], the residual parallax field after
surface alignment, or in other words the motion of objects af-
ter image registration, is an epipolar field. This is explained
by a very clear drawing reproduced in Figure 2. Let \( R_2 \) be the
ray joining a 3D point \( P \) and the center of camera 2 (denoted
\( M \)); \( R_2 \) intersects a given arbitrary surface at a 3D point de-
noted \( Q \). The projection of the ray \( R_2 \) on camera 1 includes
the projections of the underlying 3D points \( Q, P \) and \( M \) (de-
noted \( q, p, \) and \( m \) respectively). Since \( m \) is the epipole in
image 1, this means that the residual parallax vector \( \vec{q}_p \), after
registration of both images with respect to an arbitrary sur-
face, is oriented along the epipolar line \( \vec{q}_m \). Notice that \( \vec{q}_p \)
can either be directed towards or opposite to the direction of
the epipole depending on the relative position of 3D point \( P \)
and the surface. On the contrary, because true changes appear
in only one of the two images, the residual parallax vector
will not be defined. Therefore, reasoning on residual parallax
vectors, which we obtain using optical flow, provides an ef-
fective way to distinguish between true changes and parallax
effects.

2. PARALLAX EFFECTS AND EPIPOLAR
GEOMETRY

In order to be robust to viewpoint differences, an algorithm
should distinguish between real changes and parallax effects.
As demonstrated in [7], the residual parallax field after
surface alignment, or in other words the motion of objects af-
ter image registration, is an epipolar field. This is explained
by a very clear drawing reproduced in Figure 2. Let \( R_2 \) be the
ray joining a 3D point \( P \) and the center of camera 2 (denoted
\( M \)); \( R_2 \) intersects a given arbitrary surface at a 3D point de-
noted \( Q \). The projection of the ray \( R_2 \) on camera 1 includes
the projections of the underlying 3D points \( Q, P \) and \( M \) (de-
noted \( q, p, \) and \( m \) respectively). Since \( m \) is the epipole in
image 1, this means that the residual parallax vector \( \vec{q}_p \), after
registration of both images with respect to an arbitrary sur-
face, is oriented along the epipolar line \( \vec{q}_m \). Notice that \( \vec{q}_p \)
can either be directed towards or opposite to the direction of
the epipole depending on the relative position of 3D point \( P \)
and the surface. On the contrary, because true changes appear
in only one of the two images, the residual parallax vector
will not be defined. Therefore, reasoning on residual parallax
vectors, which we obtain using optical flow, provides an ef-
cfective way to distinguish between true changes and parallax
effects.

3. OVERVIEW OF OUR ALGORITHM

As discussed in the previous section, images should be
aligned with respect to an arbitrary surface. A given test
image is registered to the reference one using a homography
transformation. First, SURF keypoints are extracted from
both images, possible matches are identified using their de-
scriptors and the epipolar geometry is then estimated using
the Random Sample Consensus (RANSAC) algorithm. No-
tice that this step may be accelerated if camera parameters
are roughly known both for the reference and the test images
(which is usually the case in the context of our applications).
Afterwards, more accurate keypoint matches, based on the
epipolar geometry, together with RANSAC are again used in
order to estimate the homography transformation.

In order to estimate residual parallax vectors, we use the
optical flow algorithm described in [8], which is fast and pro-
vides a dense vector field. Experiments showed that although
most of the optical flow vectors are consistent with the epipo-
lar field (see Figure 3), some of them are spurious and are a
source of false alarms in the resulting change mask. Conse-
quently, we constrain, in our algorithm, optical flow vectors to
be collinear to the epipole direction. Then we detect changes
using a likelihood ratio test on the local matching score.

Let \( X \) be a random variable standing for a pixel into a test
image. Let \( H_0 \) denote the “null hypothesis” meaning that \( X \)
is unchanging (according to a well established ground truth);
the alternative hypothesis (denoted \( H_1 \)) stands for the con-
verse (i.e., \( X \) is changing). Now, we consider the following
likelihood ratio and its corresponding threshold:

\[
L(X) = \frac{P(\epsilon(X, \psi(X))|H_1)}{P(\epsilon(X, \psi(X))|H_0)}
\]

\[
\tau = \frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})},
\]

where \( \epsilon \) denotes an error function and \( \psi \) is the optical flow
mapping which given \( X \) in the test image, finds its unique
matching point in the reference image. \( C_{ij} \) is a cost associ-
ated with making a decision in favor of hypothesis \( H_i \) when
\( H_j \) is true. The decision as whether \( X \) is changing depends on
the sign of \( L(X) - \tau \). In order to model \( P(\epsilon(X, \psi(X)) | H_1) \)
Fig. 4. This figure describes the probability density functions of optical flow matching errors for Changed and Unchanged pixels.

Fig. 5. This figure shows the reference image (a), test image (b) with change (circle), ground truth (c) and estimated (d) change masks.

and $P(\epsilon(X, \psi(X)) \mid H_0)$, we analyze the optical flow matching error $\epsilon(X, \psi(X))$ through a population of changing and unchanging points using our dataset (see section 4). The distributions of these errors are described in Figure 4 with the corresponding models.

Figure 5 illustrates an example of change detection results obtained on a test image. As we can see, the optical flow matching score corresponds very accurately to the change mask of the ground truth. In mono-thread mode, this result was estimated in less than 3 seconds on a standard 2.4 GHz PC.

4. EVALUATION

In order to evaluate our approach quantitatively, we built the Aerial Imagery Change Detection (AICD) dataset using realistic aerial images rendered artificially\(^1\). A hundred scenes were created, with and without changes and illumination differences. Although reference and test images include illumination variations, dealing with hard shadows is out of the scope of this paper. Every scene was rendered from five different viewpoints, the last four viewpoints being more and more distant from the first one (by steps of 10 degrees), enabling the analysis of the influence of viewpoint difference on change detection performances. In addition, the true pixel-based change mask is available for each viewpoint, allowing an objective comparison of performances. This dataset will be made publicly available\(^2\) to enable benchmarking of change detection algorithms.

In order to compare detection results with respect to the ground truth, we define a precision recall criterion at the component level (instead of the pixel level). This criterion is based on intersecting connected components which is more relevant to our objective of attention focusing than a pixel-based comparison.

Figure 6 compares these precision recall performances obtained for a viewpoint difference of 10 degrees, using the method presented above with and without the epipolar geometry constraint. These curves show that constraining the optical flow algorithm using the epipolar geometry makes it possible to achieve far better performance than leaving the optical flow unconstrained, especially the number of false alarms. Indeed, unconstrained optical flow may result into wrong matches between reference and test images. Constraining the optical flow using the epipolar geometry reduces such errors and thereby the false alarm rate.

Notice that the curve obtained without the epipolar geometry constraint converges to the bottom-left corner of the diagram in Figure 6. This behavior corresponds to a scoring

---

\(^1\)We used the built-in rendering engine of the serious game Virtual Battle Station 2, edited by BISimulations.

\(^2\)The dataset will be hosted on ComputerVisionOnline.com.
As shown in Figure 6, considering the epipolar geometry constraint attenuates these effects, as it aligns the orientation of optical flow to match the true direction, which makes the results more reliable. Figure 7 shows performances obtained on two subsets of images including scenes with and without relief. The underlying ground truth (relief vs. non relief) was generated using a hard decision on the rate of outliers during the estimation of homography with RANSAC (see section 3); resulting into two subsets of comparable sizes. Figure 7 shows that the performance does not significantly decrease in the constrained case, whereas in the unconstrained one, the performance drops by 50%. These results demonstrate that the proposed approach increases robustness to the presence of relief in images, and more generally to poor preliminary registrations between images.

Finally, using the five viewpoints available for each scene, we analyze the performances of our approach with respect to viewpoint variation. Figure 8 shows a consistent gain of the constrained approach compared to the unconstrained one. Limited variations of the viewpoint angle (10 degrees in practice) degrade the performance by less than 20% for the constrained case while for the unconstrained one, the performance decreases by more than 50%. For viewpoint angles larger than 10 degrees, the two methods show similar behaviors. This may result from the presence of challenging occlusions, due to important viewpoint variations, thereby introducing many false detections.

5. CONCLUSION

In this paper, we introduced a change detection algorithm which successfully handles challenging aerial images with strong parallax effects, reliefs and illumination variations while being simple and fast. Obtained results are satisfactory both in terms of visual inspection and Precision/Recall measures. Indeed, our method demonstrates its ability to achieve good robustness to relief and viewpoint changes.

It is known that strong shadows and occluded objects are also challenging for change detection. As a future work, we will address these issues for the purpose of achieving more effective and real-time change detection in aerial videos.

6. REFERENCES