EVALUATING STITCHING QUALITY

Jani Boutellier, Olli Silvén, Lassi Korhonen
Machine Vision Group, University of Oulu, P.O. Box 4500, 90014 University of Oulu, Finland

Marius Tico
Nokia Research Center, P.O. Box 100, 33721 Tampere, Finland

Keywords: Stitching, image mosaics, panoramas, image quality.

Abstract: Until now, there has been no objective measure for the quality of mosaic images or mosaicking algorithms. To mend this shortcoming, a new method is proposed. In this new approach, the algorithm that is to be tested, acquires a set of synthetically created test images for constructing a mosaic. The synthetic images are created from a reference image that is also used as a basis in the evaluation of the image mosaic. To simulate the effects of actual photography, various camera-related distortions along with perspective warps, are applied to the computer-generated synthetic images. The proposed approach can be used to test all kinds of computer-based stitching algorithms and presents the computed mosaic quality as a single number.

1 INTRODUCTION

Image stitching is a method for combining several images into one wide-angled mosaic image. Computer-based stitching algorithms and panorama applications have been used widely for more than ten years (Davis, 1998), (Szeliski, 1994). Although it is evident that technical improvements have taken place in computer-based image stitching, there has been no objective measure for proving this trend. Subjectively, it is relatively easy to say whether a mosaic image has flaws or not (Su et al., 2004), but analyzing the situation computationally is not straightforward at all.

If we assume that we have a mosaic image and wish to evaluate it objectively, the first arising problem is usually the lack of a reference image. Even if we had a reference image of the same scene, we would generally notice that it did not have exactly the same projection as the mosaic image, therefore making pixel-wise comparison impossible. Also, it may happen that between taking the hypothetical reference image and the narrow-angled mosaic image parts, the scene might have changed somewhat, making the comparison unfit.

In this text a method is described to overcome these problems, but before going deeper into the topic, some terms need to be agreed upon. From here on, the narrow-angle images that are consumed by a stitching algorithm, are called *source images*. Also, we will call the group of source images a sequence, even if the source images are stored as separate image files. It is worth mentioning that the source images are given to the stitching algorithm in the same order as they have been created.

To be able to create a method of evaluating mosaics, we have to know what kinds of errors exist in mosaic images. Flaws that we will call *discontinuities* are caused by unsuccessful registration of source images. Apart from completely failed registration, the usual cause of these kinds of errors is the use of an inadequate registration method. For example, if the registration method of a stitching algorithm is unable to correct perspective changes of the source images to the mosaic, the mosaic will have noticeable boundaries.

Blur is a common flaw in most imaging occasions and may be caused by the imaging device or by extrinsic causes, e.g. camera motion. In mosaicking, blur can also be caused by inadequate source image blending.

Object clipping happens when the location of an object changes in the view of the camera between the source image captures. In a practical image mosaic, a common clipped object is for example a pedestrian.

The final mosaic flaw introduced here can only be

Boutellier J., Silvén O., Korhonen L. and Tico M. (2007).
EVALUATING STITCHING QUALITY.
In *Proceedings of the Second International Conference on Computer Vision Theory and Applications - IFP/IA*, pages 10-17

Table 1: Common flaws in image mosaics.

Type of Error	Cause
Discontinuity	Failed or inadequate source image registration
Blur	Shooting conditions, unfit blending, lens distortions
Object Clipping	Moving object in source image ignored
Intensity Change	Color balancing between mosaic parts

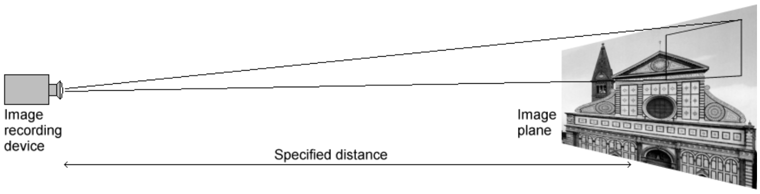

Figure 2: The simulated imaging environment.

where the camera is at some fixed distance from the reference image that appears as a plane in 3D-space. (See Figure 2). As a result, perspective distortions appear in the source images. In literature this is called the *pinhole camera model* (Jähne, 1997).

Images that are normally delivered to a stitching algorithm for mosaicking have several kinds of distortions caused by the camera and shooting conditions. Real-world camera lenses cause vignetting and radial distortions to the image (Jähne, 1997). If the user takes pictures freehand, the camera shakes and may cause motion blur to the pictures. Also, it is common that the camera is slowly rotated from one frame to another when shooting many pictures of the same scene. Finally, cameras tend to adapt to lighting conditions by changing exposure times according to the brightness of the view that is shown through the camera lens.

The frames are created by panning the simulated camera view over the reference image in a zig-zag pattern and by taking shots with a nearly constant interval (see Figure 3). The camera jitter is modeled by random vertical and horizontal deviations from the sweeping pattern along with gradually changing camera rotation. Motion blur caused by camera shaking is simulated by filtering the source image with a point-spread function consisting of a line with random length and direction.

Camera lens vignetting is implemented by multiplying the source image with a two-dimensional mask, that causes the image intensity to dim slightly as a function of the distance from the image center. Radial distortions are created by a simple function that is given in equation 1.

$$d_n = d + kd^3, \qquad (1)$$

where

d is the distance from the image center,
d_n is the new distance from the center and
k is the distortion strength parameter.

The radial distortion is applied by calculating a new distance for every pixel from the image center in the source image. The result of this warp is a barrel distortion if the constant k is positive. Finally, the simulated differences in exposure time are applied to the source image by normalizing the mean of the image to a constant value. Figure 4 shows the effect of each step in this simulated imaging process.

The source image sequence is recorded as an uncompressed video clip by default, but can of course be converted to other forms depending on the required input type of the algorithm that is chosen for testing.

3.2 Mosaic Image Registration

The stitched mosaic image and the reference image have different projections because the stitching software has had to fit together the source images that contain non-linear distortions. The mosaic image has to be registered to the coordinates of the reference image to make the comparison eligible.

For this purpose we selected a SIFT-based (Lowe, 2004) feature detection and -matching algorithm[1], that produced around one thousand matching feature points for each image pair in our tests. An initial registration estimate is calculated by a 12-parameter polynomial model, after which definite outliers are removed from the feature point set. The final registration is made by the unwarpJ -algorithm (Sorzano et al., 2005) that is based on a B-spline deformation model. It is evident that the success of image registration is a most important factor to ensure the eligibility

[1] http://www.cs.ubc.ca/ lowe/keypoints/

of the quality measurement. According to the conducted tests, the accuracy and robustness of unwarpJ are suitable for the purpose.

An example of the registration process can be seen in Figure 6. The topmost image is a mosaic image created by a stitching algorithm. The middle image in Figure 6 shows the registered version of the mosaic above, and the image in the bottom shows the corresponding *similarity map* (See next subsection). Notice how slight stitching errors have prevented flawless matching in the bottom-right corner of the image.

3.3 Similarity Calculation

Our method uses the recent approach of Wang (Wang et al., 2004) to estimate the quality of the registered mosaic. The approach of Wang estimates the similarity of two images and gives a single similarity index value that tells how much alike the two images are. This method fits very well to the requirements of mosaic evaluation, since it pays attention on distortions that are clearly visible for the human vision system. This includes blurring and structural changes that are common problems in mosaics. Wang's method does not penalize for slight changes in the image intensity.

This arrangement will notice blurring and discontinuity -flaws in the mosaic images that are created from test videos. With the current test setup it is not possible to simulate situations that would cause object clipping.

4 PRACTICAL TESTS

We used three different stitching algorithms to test the functionality of our test method: Autostitch (Brown and Lowe, 2003), Surveillance Stitcher (Heikkilä and Pietikäinen, 2005) and a still unpublished algorithm called Mobile Stitcher. The algorithms were tested by three different video sequences that were created from the images shown in Figure 5.

Figure 3: Camera motion pattern over a reference image and a quadrangle depicting the area included in an arbitrary source image.

Each algorithm was tested with the three sequences. Results of subjective mosaic evaluation and quality indexes provided by our algorithm are visible in Table 2. A detailed analysis concerning one of the sequences can be found in the caption of Figure 7. The other results are only shown as pictures due to space constraints.

We can notice from the results that the acquired quality indexes are not comparable from one sequence to another. The focus of this testing was not to sort the tested algorithms to some order of quality, but to simply show what kinds of results can be achieved with our testing method. When the similarity indexes were calculated, 50 pixels from each image border were omitted, since the non-linear registration algorithm was often a bit inaccurate near image borders.

UBC Autostitch acquired the best results from each test, which can also be detected visually, since the results are practically absent of discontinuities. The Surveillance Stitcher acquired second best results, although most of its results had slight discontinuities. The Mobile Stitcher performed worst in these tests, which is easily explained by the fact that the algorithm is the only one of the three that uses an area-based registration (Zitová and Flusser, 2003) method and thus is unable to correct perspective distortions.

Figure 8 shows a mosaic that was created by Autostitch along with some modified versions of the mosaic. The figure depicts how different kinds of mosaicking errors affect the similarity map and the numerical quality of the mosaic. A more detailed explanation can be found in the caption of Figure 8.

A slight setback in the testing was to notice that the currently used registration method proved to be inaccurate in one occasion. In Figure 8 each similarity map indicates that something is wrong in the right end of the building. However, visual inspection reveals no problems. The reason behind the indicated difference is effectively a misalignment of a few pixels. This should obviously be corrected in the future.

5 CONCLUSION

We have presented a novel way to measure the performance of stitching algorithms. The method is directly applicable to computer-based algorithms that can create mosaic images from source image sequences.

The method could be improved by using depth-varying 3D models, which would test the algorithms' abilities to cope with effects of occlusion and paral-

Table 2: Mosaic quality values and visually observed distortions of test mosaics. The numbers in braces indicate which figure displays the corresponding result, if it is shown.

Algorithm	Pattern	Facade	Graffiti
UBC Autostitch	0.81 (7), blur, slight discont.	0.89 (8), blur	0.76, blur
Mobile stitcher	0.80 (7), discontinuities	0.68, discontinuities	0.65, discontinuities
Surveillance stitcher	0.75 (7), blur, discont.	0.86, blur, slight discont.	0.72 (6), blur, discont.

lax. Also, the presence of moving objects could be simulated in future versions of the testing method.

REFERENCES

Bors, A. G., Puech, W., Pitas, I., and Chassery, J.-M. (1997). Perspective distortion analysis for mosaicing images painted on cylindrical surfaces. In *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages 3049–3052.

Brown, M. and Lowe, D. G. (2003). Recognising panoramas. In *Proceedings of the 9th International Conference on Computer Vision*, volume 2, pages 1218–1225.

Davis, J. (1998). Mosaics of scenes with moving objects. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 354–360.

Feldman, D. and Zomet, A. (2004). Generating mosaics with minimum distortions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pages 163–170.

Heikkilä, M. and Pietikäinen, M. (2005). An image mosaicing module for wide-area surveillance. In *Proceedings of the ACM Workshop on Video Surveillance and Sensor Networks*, pages 11–18.

Jähne, B. (1997). *Digital Image Processing, Concepts, Algorithms, and Scientific Applications*. Springer-Verlag, Berlin, 4th edition.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110.

Marzotto, R., Fusiello, A., and Murino, V. (2004). High resolution video mosaicing with global alignment. In *Proceedings of the IEEE Conference of Computer Vision and Pattern Recognition*, volume 1, pages 692–698.

Sorzano, C. O., Thevenaz, P., and Unser, M. (2005). Elastic registration of biological images using vector-spline regularization. *IEEE Transactions on Biomedical Engineering*, 52(4):652–663.

Su, M.-S., Hwang, W.-L., and Cheng, K.-Y. (2004). Analysis on multiresolution mosaic images. *IEEE Transactions on Image Processing*, 13(7):952–959.

Swaminathan, R., Grossberg, M. D., and Nayar, S. K. (2003). A perspective on distortions. In *Proceedings of the IEEE Conference of Computer Vision and Pattern Recognition*, volume 2, pages 594–601.

Szeliski, R. (1994). *Image Mosaicing for Tele-Reality Applications*. DEC Cambridge Research Lab Technical Report 94/2.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612.

Zitová, B. and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000.