LaTeX Author Guidelines for CVPR Proceedings

First Author
Institution1
Institution1 address
firstauthor@i1.org

Second Author
Institution2
First line of institution2 address
secondauthor@i2.org

Abstract

The ABSTRACT is to be in fully-justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word "Abstract" as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. Leave two blank lines after the Abstract, then begin the main text. Look at previous CVPR abstracts to get a feel for style and length.

1. Introduction

Reconstructing a 3D object from input images is an essential task in the field of computer vision. Recent advances in photo databases have led to an increase in the use of 3D reconstruction. While there have been many recent advances in 3D reconstruction, most new techniques still suffer from the same problems which plagued original methods. First, many approaches require a camera calibration for each image used in the 3D reconstruction. This calibration is necessary to obtain the intrinsic and extrinsic parameters of the camera which are used to project points from 2D images into 3D space. Performing calibrations for each image is impractical for new approaches that emphasize using as many images as possible for the reconstruction. The other setback relates to user input. Many approaches require users to pick points on the input images that specify either the object to reconstruct or the planes used to perform the reconstruction. While user input may improve the reconstruction as it improves depth detection and crops the image so only necessary points are used, it does not encourage using many input images. Several techniques overcome these setbacks by performing reconstruction based on uncalibrated images obtained by the same, freely moving camera. Camera calibration are performed once and the resulting parameters are used for the reconstruction across all input images. Capturing pairs of feature points from the input images is an important step in the 3D reconstruction process. There are several feature point detection algorithms which are currently used by many reconstruction techniques, though each has its own setbacks. Two popular approaches are Scale Invariant Feature Transform (SIFT) and corner extraction algorithms such as Harris corner detection. Each approach yields pairs of matched feature points across the input images. In this project, we develop a new technique which reconstructs a 3D object from two input images captured by a freely moving camera. We then combine SIFT and Harris corner detection to obtain the relevant pairs of feature points from the two input images. Prior to using these matching points, we use the Random Sample Consensus (RANSAC) algorithm to remove outlier matches. We calculate the intrinsic parameters of the camera used to take the input images with Zhangs method. Using the intrinsic parameters of the camera and the matched feature points from the input images, we are able to obtain a 3D reconstruction for the primary object in the image.

2. Related Work

3. Overview

In this section, we give an overview of our algorithm, and in later sections, we describe the different components in more details.

In this section, we give an overview of our algorithm, which we will describe in more detail in later sections.

1. We find the intrinsic parameters of our camera using Zhang's method.

2. Using SIFT and Harris corner detection, we identify pairs of features points on both images, and we use RANSAC to find the inliers so that we can generate a representative fundamental matrix.

3. With the intrinsic parameters of our camera and the fundamental matrix, we can derive the essential matrix.

4. The essential matrix is used to derive the projection matrices of the images which each comprise of a rota-

tion matrix and translation vector. With this information, we can reconstruct the 3d image using triangulation.

5. With this information, we can reconstruct the 3d image using triangulation.

4. Camera Calibration

To calibrate our camera and obtain its intrinsic parameters, we use Zhangs method. Zhangs method is a technique that uses several images to derive a cameras focal length, aspect ratio, and principal points. Whilethere are many other approaches to performing this essential task, Zhangs method is more flexible and robust. The two traditional approaches are often categorized as either photogrammetric or self-calibration. Photogrammetric methods require using a 3D object whose 3D coordinates are precisely known. Taking multiple images of this 3D object lets one infer the intrinsic parameters of the camera as they can be derived from the difference between the actual coordinates of the object and what is seen across the images. However, the apparatus necessary to perform this type of calibration is expensive. Self-calibration, while less costly, is not reliable as there are often not enough known points to estimate all the necessary parameters. Self-calibration requires moving the camerain a static setup and performing a feature points matching across the taken images. Self-calibration also uses constraints pertaining to the rigidity of the object considered.

Zhangs method is a cross between photogrammetric and self-calibration techniques. Zhangs method requires one to construct a pattern on a paper and attach it to a planar surface. Several images of the pattern, from different angles, are then taken. As long as the camera or the pattern is stationary, the movement is not restricted. Using an understanding of the geometry of the designed pattern, constraints on the intrinsic parameters arise from each view. Using all of these constraints, a set of intrinsic parameters which satisfy all the considered views can be inferred. Of course, using more views will yield intrinsic parameters closer to their actual values. Thus, Zhangs method incorporates understanding 3D coordinates of the pattern (photogrammetric) and using multiple views to set constraints on the intrinsic parameters (self-calibration), but it only considers one plane and a user-designed pattern. In their paper, Zhang et. al. show that their calibration technique is both flexible, robust, and accurate.

A cameras intrinsic matrix is defined as:

$$A = \begin{bmatrix} \alpha & \gamma & u0 \\ 0 & \beta & v0 \\ 0 & 0 & 1 \end{bmatrix}$$

where (u0, v0) is the principal point, α and β are image scale factors, and γ is a parameter which describes the skew in the image axes. We use the pinhole camera model which dictates that a 3D point, M, is related to its image projection, m, by:

$$sm = A \begin{bmatrix} R & t \end{bmatrix} M \tag{1}$$

where R is the rotation matrix and t is the translation vector of the considered image. For each view used, we estimate a homography matrix:

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = \lambda A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \tag{2}$$

where r_1 and r_2 are the components of the rotation matrix, R, and lambda is a scaling factor. Since r_1 and r_2 are orthonormal, we can derive the following two constraints:

$$h_1^T A^{-T} A^{-1} h_2 = 0 \tag{3}$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \tag{4}$$

The approximation of the cameras intrinsic parameters is improved using a maximum-likelihood estimation. Zhang's method also extracts the cameras extrinsic parameters for the given input images, but we do not make use of this data as we are reconstructing an object separate from the considered planar pattern.

5. Fundamental and Essential Matrix

The fundamental matrix contains information about the intrinsic and extrinsic parameters of the camera.

6. 3D reconstruction

We use the fundamental and essential matrix to reconstruct the 3d image.

7. Results

Here we show some of our reconstructions. We used a camera with the following intrinsic parameters.

8. Conclusion

This project describes a flexible and practical 3D-reconstruction technique. Rather than having to perform a camera calibration for each input image used in the reconstruction, as many previous techniques have required, we only perform the calibration once using Zhangs method. Thus, this approach scales well with the number of input images used, as long as they are all taken with the same camera. We also combine several common feature point extraction methods to enhance the detail in the final 3D reconstruction. Using both SIFT and Harris corner detection, we are able to get enough pairs of matching points to recover all eight unknowns in the fundamental matrix while incorporating all the necessary feature points to capture depth

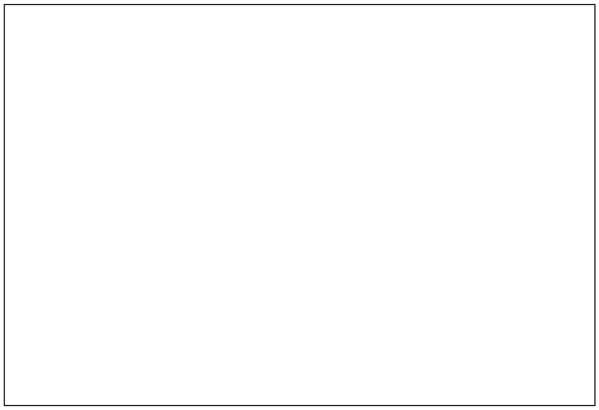

Figure 1. Example of caption. It is set in Roman so that mathematics (always set in Roman: $B \sin A = A \sin B$) may be included without an ugly clash.

and structural details in the reconstructed object. The benefits of including Harris corner detection are highlighted in the results we provided. Both edges and corners are much more pronounced in the reconstructed images using corner detection than in those when we used SIFT alone.

8.1. Miscellaneous

Compare the following:

`$conf_a$`	$conf_a$
`$\mathit{conf}_a$`	conf_a

See The TeXbook, p165.

The space after *e.g.*, meaning "for example", should not be a sentence-ending space. So *e.g.* is correct, *e.g.* is not. The provided `\eg` macro takes care of this.

When citing a multi-author paper, you may save space by using "et alia", shortened to "*et al.*" (not "*et. al.*" as "*et*" is a complete word.) However, use it only when there are three or more authors. Thus, the following is correct: " Frobnication has been trendy lately. It was introduced by Alpher [**?**], and subsequently developed by Alpher and Fotheringham-Smythe [**?**], and Alpher *et al*. [**?**]."

This is incorrect: "... subsequently developed by Alpher *et al*. [**?**] ..." because reference [**?**] has just two authors. If you use the `\etal` macro provided, then you need not worry about double periods when used at the end of a sentence as in Alpher *et al*.

For this citation style, keep multiple citations in numerical (not chronological) order, so prefer [**?**, **?**, **?**] to [**?**, **?**, **?**].

8.2. Margins and page numbering

All printed material, including text, illustrations, and charts, must be kept within a print area 6-7/8 inches (17.5 cm) wide by 8-7/8 inches (22.54 cm) high. Page numbers should be in footer with page numbers, centered and .75 inches from the bottom of the page and make it start at the correct page number rather than the 4321 in the example. To do this fine the line (around line 23)

```
%\ifcvprfinal\pagestyle{empty}\fi
\setcounter{page}{4321}
```

where the number 4321 is your assigned starting page.

Make sure the first page is numbered by commenting out the first page being empty on line 46

```
%\thispagestyle{empty}
```

8.3. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [**?**]. Where appropriate, include the name(s) of editors of referenced books.

Method	Frobnability
Theirs	Frumpy
Yours	Frobbly
Ours	Makes one's heart Frob

Table 1. Results. Ours is better.

8.4. Illustrations, graphs, and photographs

All graphics should be centered. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text, and choose line widths which render effectively in print. Many readers (and reviewers), even of an electronic copy, will choose to print your paper in order to read it. You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic.

When placing figures in LaTeX, it's almost always best to use `\includegraphics`, and to specify the figure width as a multiple of the line width as in the example below

```
\usepackage[dvips]{graphicx} ...
\includegraphics[width=0.8\linewidth]
                {myfile.eps}
```

8.5. Color

Color is valuable, and will be visible to readers of the electronic copy. However ensure that, when printed on a monochrome printer, no important information is lost by the conversion to grayscale.

9. Final copy

You must include your signed IEEE copyright release form when you submit your finished paper. We MUST have this form before your paper can be published in the proceedings.

Figure 2. Example of a short caption, which should be centered.