

L^AT_EX Author Guidelines for ICCV 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 ICCV abstracts to get a feel for style and length.

1. Introduction

3D face reconstruction with fine scale surface details has a wide range of application. Though there are highly robust methods of reconstructing high-quality 3D faces with fine geometry details using stereo photography [] and the photometric stereo method [], those techniques usually need expensive and bulky capture systems. Therefore, it is still challenging and worthwhile to explore approaches to construct fine detailed 3D faces merely from limited data source.

This paper focuses on reconstruction of 3D faces from a single image.

With input limited to one single image, one of the most common approaches is to first build a 3D proxy derived from a 3D morphable model(3DMM[]) and then refine the proxy by deforming the mesh so that the proxy fits to the original image.

However, 3D morphable models themselves use a small number of parameters, which lead to over-smooth reconstructed surfaces and the lack of facial details like wrinkles and dimples. [Shape from shading], the special case of photometric stereo where only a single image is used, is the traditional method used to handle loss of details. It is able to recover high frequency details by estimating surface normal according to reflection of light in the image. However, the method may be error-prone and sensitive to noise.

Recently, deep learning based methods have achieved impressive progress in reconstruction of subtle facial de-

tails and expressions. [pix2vertex] directly learns a depth map and a position correspondence map from face images, and obtains 3D facial mesh by using non-rigid transformation; [Photo-Realistic...] predicts emotion info from input images via the proposed EmotionNet, and uses that info to determine expression parameters of 3DMM during proxy face geometry generation. On facial detail synthesis, they achieve high realistic facial details by predicting a displacement map via a semi-supervised learning procedure; [DF2Net/others] the recent state-of-the-art methods are based on the coarse-to-fine framework, where the so-called coarse and fine networks are cascaded in an end-to-end fashion. The coarse network handles reconstruction of coarse face geometry like the 3DMM’s model fitting process, while the fine network add facial details.

All methods above require a training dataset large and generalized enough which describes fine-scale facial geometry. At present, plenty of high-resolution face images are available on the Internet. However, large-scale datasets providing connections between facial images and their high-frequency geometry details are still lacking.

In order to alleviate this problem, [photo-realistic] manually scans a relatively small amount of 3D faces with high quality geometry details, and exploits 163K facial images in the wild via an additional unsupervised learning procedure. In [DF2Net], real images containing facial details are also collected for training, but different from [photo-realistic], they run existing reconstruction algorithm on those images and regard the results as ground-truth geometry details, so that supervised learning can be employed.

Compared to employing unsupervised learning approaches or using 3D meshes obtained by existing reconstruction algorithms, it may be more effective to synthesize training data directly. As far as we know, synthesis methods are currently only used to generate geometry of whole faces with varying expressions, poses and face shapes, and there are currently no large-scale synthesized dataset that contains mid- and high-frequency facial geometry details.

In this paper, we propose a novel synthesis method for skin detail generation.

2. Formatting your paper

All text must be in a two-column format. The total allowable width of the text area is $6\frac{7}{8}$ inches (17.5 cm) wide by $8\frac{7}{8}$ inches (22.54 cm) high. Columns are to be $3\frac{1}{4}$ inches (8.25 cm) wide, with a $\frac{5}{16}$ inch (0.8 cm) space between them. The main title (on the first page) should begin 1.0 inch (2.54 cm) from the top edge of the page. The second and following pages should begin 1.0 inch (2.54 cm) from the top edge. On all pages, the bottom margin should be 1-1/8 inches (2.86 cm) from the bottom edge of the page for 8.5×11 -inch paper; for A4 paper, approximately 1-5/8 inches (4.13 cm) from the bottom edge of the page.

2.1. Margins and page numbering

All printed material, including text, illustrations, and charts, must be kept within a print area $6\frac{7}{8}$ inches (17.5 cm) wide by $8\frac{7}{8}$ inches (22.54 cm) high.

Page numbers should be included for review submissions but not for the final paper. Review submissions papers should have page numbers in the footer with numbers centered and .75 inches (1.905 cm) from the bottom of the page and start on the first page with the number 1.

Page numbers will be added by the publisher to all camera-ready papers prior to including them in the proceedings and before submitting the papers to IEEE Xplore. As such, your camera-ready submission should not include any page numbers. Page numbers should automatically be removed by uncommenting (if it's not already) the line

```
% \iccvfinalcopy
```

near the beginning of the .tex file.

2.2. Type-style and fonts

Wherever Times is specified, Times Roman may also be used. If neither is available on your word processor, please use the font closest in appearance to Times to which you have access.

MAIN TITLE. Center the title $1\frac{3}{8}$ inches (3.49 cm) from the top edge of the first page. The title should be in Times 14-point, boldface type. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs; do not capitalize articles, coordinate conjunctions, or prepositions (unless the title begins with such a word). Leave two blank lines after the title.

AUTHOR NAME(s) and **AFFILIATION(s)** are to be centered beneath the title and printed in Times 12-point, non-boldface type. This information is to be followed by two blank lines.

The **ABSTRACT** and **MAIN TEXT** are to be in a two-column format.

MAIN TEXT. Type main text in 10-point Times, single-spaced. Do NOT use double-spacing. All paragraphs

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

Table 1. Results. Ours is better.

should be indented 1 pica (approx. 1/6 inch or 0.422 cm). Make sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

Figure and table captions should be 9-point Roman type as in Figures ?? and ??. Short captions should be centered. Callouts should be 9-point Helvetica, non-boldface type. Initially capitalize only the first word of section titles and first-, second-, and third-order headings.

FIRST-ORDER HEADINGS. (For example, **1. Introduction**) should be Times 12-point boldface, initially capitalized, flush left, with one blank line before, and one blank line after.

SECOND-ORDER HEADINGS. (For example, **1.1. Database elements**) should be Times 11-point boldface, initially capitalized, flush left, with one blank line before, and one after. If you require a third-order heading (we discourage it), use 10-point Times, boldface, initially capitalized, flush left, preceded by one blank line, followed by a period and your text on the same line.

2.3. Footnotes

Please use footnotes¹ sparingly. Indeed, try to avoid footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence). If you wish to use a footnote, place it at the bottom of the column on the page on which it is referenced. Use Times 8-point type, single-spaced.

2.4. 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 [1]. Where appropriate, include the name(s) of editors of referenced books.

2.5. 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.

¹This is what a footnote looks like. It often distracts the reader from the main flow of the argument.

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}
```

2.6. Color

Please refer to the author guidelines on the ICCV 2019 web page for a discussion of the use of color in your document.

3. 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.

References

- [1] Full Author Name. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as additional material `fg324.pdf`.