LaTeX Author Guidelines for CVPR Proceedings

First Author
Institution1
Institution1 address
firstauthor@i1.org

Second Author
Institution2
First line of institution2 address
http://www.author.org/~second

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

Please follow the steps outlined below when submitting your manuscript to the IEEE Computer Society Press. This style guide now has several important modifications (for example, you are no longer warned against the use of sticky tape to attach your artwork to the paper), so all authors should read this new version.

1.1. Language

All manuscripts must be in English.

1.2. Dual submission

By submitting a manuscript to CVPR, the authors assert that it has not been previously published in substantially similar form. Furthermore, no paper which contains significant overlap with the contributions of this paper either has been or will be submitted during the CVPR 2009 review period to **either a journal** or a conference (including CVPR 2009). **Note that this is a strengthening of previous CVPR policy,** and papers violating this condition will be rejected.

If there are papers that may appear to the reviewers to violate this condition, then it is your responsibility to: (1) cite these papers (preserving anonymity as described in Section 1.6 below), (2) argue in the body of your paper why your CVPR paper is non-trivially different from these concurrent submissions, and (3) include anonymized versions of those papers in the supplemental material.

1.3. Paper length

Consult the call for papers for page-length limits. Over-length papers will simply not be reviewed. This includes papers where the margins and formatting are deemed to have been significantly altered from those laid down by this style guide. Note that this LaTeX guide already sets figure captions and references in a smaller font. The reason such papers will not be reviewed is that there is no provision for supervised revisions of manuscripts. The reviewing process cannot determine the suitability of the paper for presentation in eight pages if it is reviewed in eleven.

1.4. The ruler

The LaTeX style defines a printed ruler which should be present in the version submitted for review. The ruler is provided in order that reviewers may comment on particular lines in the paper without circumlocution. If you are preparing a document using a non-LaTeX document preparation system, please arrange for an equivalent ruler to appear on the final output pages. The presence or absence of the ruler should not change the appearance of any other content on the page. The camera ready copy should not contain a ruler. (LaTeX users may uncomment the \cvprfinalcopy command in the document preamble.) Reviewers: note that the ruler measurements do not align well with lines in the paper — this turns out to be very difficult to do well when the paper contains many figures and equations, and, when done, looks ugly. Just use fractional references (e.g. this line is 095.5), although in most cases one would expect that the approximate location will be adequate.

1.5. Mathematics

Please number all of your sections and displayed equations. It is important for readers to be able to refer to any particular equation. Just because you didn't refer to it in

the text doesn't mean some future reader might not need to refer to it. It is cumbersome to have to use circumlocutions like "the equation second from the top of page 3 column 1". (Note that the ruler will not be present in the final copy, so is not an alternative to equation numbers). All authors will benefit from reading Mermin's description of how to write mathematics: http://www.cvpr.org/doc/mermin.pdf.

1.6. Blind review

Many authors misunderstand the concept of anonymizing for blind review. Blind review does not mean that one must remove citations to one's own work—in fact it is often impossible to review a paper unless the previous citations are known and available.

Blind review means that you do not use the words "my" or "our" when citing previous work. That is all. (But see below for techreports)

Saying "this builds on the work of Lucy Smith [1]" does not say that you are Lucy Smith, it says that you are building on her work. If you are Smith and Jones, do not say "as we show in [7]", say "as Smith and Jones show in [7]" and at the end of the paper, include reference 7 as you would any other cited work.

An example of a bad paper:

> An analysis of the frobnicatable foo filter.
>
> In this paper we present a performance analysis of our previous paper [1], and show it to be inferior to all previously known methods. Why the previous paper was accepted without this analysis is beyond me.
>
> [1] Removed for blind review

An example of an excellent paper:

> An analysis of the frobnicatable foo filter.
>
> In this paper we present a performance analysis of the paper of Smith *et al*. [1], and show it to be inferior to all previously known methods. Why the previous paper was accepted without this analysis is beyond me.
>
> [1] Smith, L and Jones, C. "The frobnicatable foo filter, a fundamental contribution to human knowledge". Nature 381(12), 1-213.

If you are making a submission to another conference at the same time, which covers similar or overlapping material, you may need to refer to that submission in order to explain the differences, just as you would if you had previously published related work. In such cases, include the anonymized parallel submission [4] as additional material and cite it as

[1] Authors. "The frobnicatable foo filter", ECCV 2006 Submission ID 324, Supplied as additional material eccv06.pdf.

Finally, you may feel you need to tell the reader that more details can be found elsewhere, and refer them to a technical report. For conference submissions, the paper must stand on its own, and not *require* the reviewer to go to a techreport for further details. Thus, you may say in the body of the paper "further details may be found in [5]". Then submit the techreport as additional material. Again, you may not assume the reviewers will read this material.

Sometimes your paper is about a problem which you tested using a tool which is widely known to be restricted to a single institution. For example, let's say it's 1969, you have solved a key problem on the Apollo lander, and you believe that the CVPR70 audience would like to hear about your solution. The work is a development of your celebrated 1968 paper entitled "Zero-g frobnication: How being the only people in the world with access to the Apollo lander source code makes us a wow at parties", by Zeus *et al*.

You can handle this paper like any other. Don't write "We show how to improve our previous work [Anonymous, 1968]. This time we tested the algorithm on a lunar lander [name of lander removed for blind review]". That would be silly, and would immediately identify the authors. Instead write the following:

> We describe a system for zero-g frobnication. This system is new because it handles the following cases: A, B. Previous systems [Zeus et al. 1968] didn't handle case B properly. Ours handles it by including a foo term in the bar integral.
>
> ...
>
> The proposed system was integrated with the Apollo lunar lander, and went all the way to the moon, don't you know. It displayed the following behaviours which show how well we solved cases A and B: ...

As you can see, the above text follows standard scientific convention, reads better than the first version, and does not explicitly name you as the authors. A reviewer might think it likely that the new paper was written by Zeus *et al*., but cannot make any decision based on that guess. He or she would have to be sure that no other authors could have been contracted to solve problem B.

FAQ: Are acknowledgements OK? No. Leave them for the final copy.

1.7. Miscellaneous

Compare the following:

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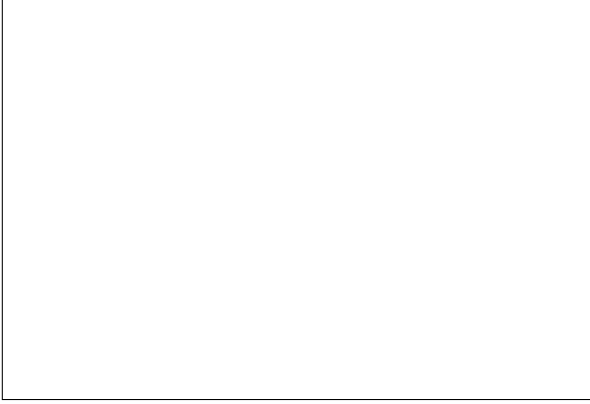$conf_a$ and conf_a

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.

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 [3], and subsequently developed by Alpher and Fotheringham-Smythe [1], and Alpher *et al.* [2]."

This is incorrect: "... subsequently developed by Alpher *et al.* [1] ..." because reference [1] 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 [1, 3, 4] to [3, 1, 4].

Please do not delete the trash contents now. It is useful to make our paper more formatted. I will delete them before submission. Thank you

2. Introduction

Introduction here

3. Related Work

Related work here.

4. Data Set

Based on our best knowledge, there is no standard high resolution image data set available on the web even though there are low resolution image data set MIT and IN-RIA. So we create our own original high resolution data set which costs us a tons of time. We take the photos in different places at Stanford University and then manually crop them into standard size and resolution images. They are all reserved as research purpose only. we keep all the rights on this data set. The finalized classification images are standard 2000*1200 resolution, in which there are about 462/596 Training Positive and Negative Images and 462/596 Testing Positive and Negative Images. There are also 829 detection images with higher resolution 3000*4000. The followings are some samples selected from our own data set:

Figure here.

In the classification images, there are generally a single standing upright person with different poses, backgrounds, scales, illuminations and occlusions marked as Positive and no person with different backgrounds and illuminations marked as Negative. In the detection images, many people with different sizes and scales appear in the images. After we build our algorithm based on classification images, we detect multi people in detection images using different windows and parameters. In order to make data set standard, we named our original data set "SUHR" short for Stanford University High Resolution data set.

5. Rough Human Classifier

In our framework, first step is to do rough human classification to get an initial results on SUHR. Our method is based on [1] but with some modifications. There are the details steps as described below.

5.1. Grayscaled, Normalized, Gamma Transformation

SUHR is colored images with different conditions but our framework is based on grayscaled clear images. We transform input images into grayscaled images and normalize images. After that we make Gamma Transformation on the images with parameter $c = 0.9$ and $\gamma = 0.8$. This standard preprocessing will ensure final performance.

5.2. Gradient, Spatial and Orientation

We calculate gradient images using mask [-1, 0, 1] in both x-direction and y-direction. From gradient images, we calculate orientation images $\varphi = \arctan(\frac{dy}{dx})$. Based on gradient and orientation images we calculate Histogram of Gradient (HoG) by cell and block. Cell is a unit to calculate the HoG and block is a group of cells to get the block descriptors. In order to get detailed information, overlapped block with by 2 cells shared by neighbored blocks is used. So each cell will be calculated four times by four different blocks. We calculate HoG on the overlapped square cell and block. In our own data set, we calculate HoG on square
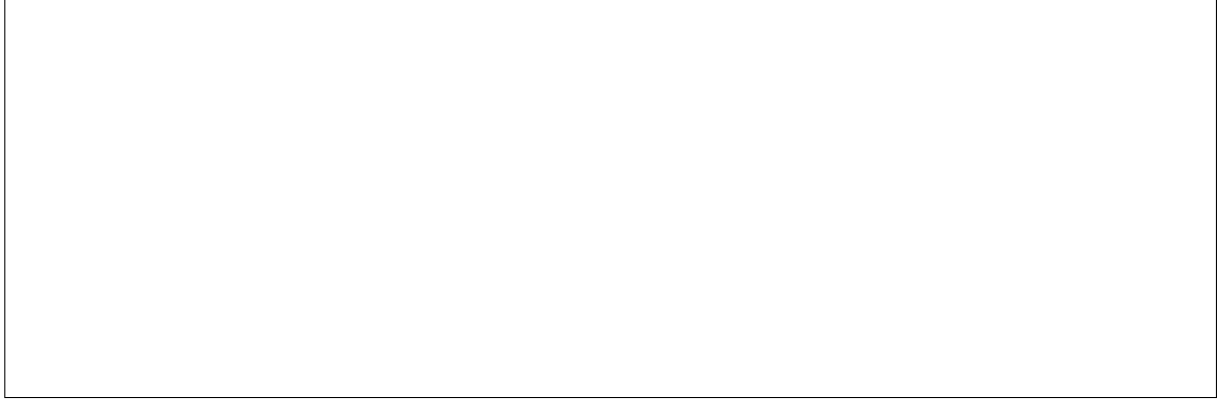
Figure 2. Example of a short caption, which should be centered.

cells, add gradients by different magnitude weights, transform angles into 9 bins ranged equally in $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and finally combine them in blocks. For high resolution images, we take the best parameters cell = 100 and block = 2 * 2 cells which will be proved effective in later section. It is useful to get weighted values using gaussian spatial window on each cell before accumulating orientation weights into cells and blocks. We got 9 bins on each cell and 9*4 descriptors on each block. Generally speaking, on high resolution images 2000*1200 we got $(\frac{2000}{100}-1)*(\frac{1200}{100}-1)*36 = 7524$ length block descriptor.

5.3. Block Descriptor Normalization

Since one block equals 2*2 cells, there maybe different illumination and background conditions so we normalize block descriptor using $L2 - norm : v \longrightarrow \frac{v}{\sqrt{((\|v\|)_2)^2+\epsilon^2}}$. The normalized block descriptor will decrease illumination and background affects and is better to represent images.

5.4. Classifier

We use linear kernel SVMlight to train our classifier based on training set and get classification prediction on testing set. It is very effective and has good performance on high dimension features.

6. Automatic Threshold Selection

After rough human classification, we get the rough prediction on our data set. Even though rough classification could get good prediction more than 90% accuracy but in order to further improve classification results we need to select bad-performance prediction to make detailed classification. In our framework, we create a novel method to automatically select this threshold by training set. Based on the prediction results on training set, we try different threshold from [-0.1, 0.1] to [-0.9, 0.9] with 0.01 step. On each selected set, we reclassify the selected set and compare final

accuracy, precision and recall. From the threshold pool we select the best threshold ats according to best performance on accuracy, precision and recall.

After afs is determined, we get selected set and begin further processing. Since the people are mostly standing upright and then head, torso and leg are relative fixed compared to image center. So in 2000*1200 images with top-left as original, we choose [] as head area, [] as torso area and [] as leg area which are all selected by average of manually repetitive pick over hundreds of images. In head area we continue to build Hair, Skin, Face classifiers and torso, leg classifiers. Hair, Skin and Face classifiers follow the same pipeline while torso and leg classifiers follow the same pipeline, all of which are described as follows.

7. Hair, Skin, Face Weak Classifier

In Hair, Skin and Face classifier, we use such framework: Harris Laplace Key Points Detector, RIFT Descriptor, K-means clustering, Building Dictionary, Data Set Representation on Dictionary, SVMlight Classification and finally get our Hair, Skin and Face classifiers.

Figure here is the diagram of pipeline.

Some of Hair, Skin and Face sample images are shown here:

Figure to show sample images.

7.1. Scale Adapted Harris Detector

According to [2], Harris Detector is based on second moment matrix. This second moment matrix is defined as:

$$M = \mu(x, \sigma_I, \sigma_D) = \begin{bmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{bmatrix}$$
$$= \sigma_D^2 g(\sigma_I) * \begin{bmatrix} L_x^2(x, \sigma_D) & L_x L_y(x, \sigma_D) \\ L_x L_y(x, \sigma_D) & L_y^2(x, \sigma_D) \end{bmatrix}$$

where σ_I is the integration scale, σ_D is the differentiation scale and L_a is the derivative computed in the a direction. This Matrix is useful to extract significant curvature

points with significant changes in the orthogonal directions i.e. corners, junctions etc. Such points are stable in lighting conditions and are representative of an image. Harris detector is based on this matrix as follows:

$cornerness = det(M) - \alpha * trace^2(M)$

Local maxima of cornerness determines the Harris Corner interest points.

7.2. Automatic Scale Selection

Automatic Scale Selection is to select the characteristic scale of a local structure for which a given function attains an extremum over scales. The selected scale measures the scale at which there is maximum similarity between feature detection operator and local image structures. This scale selection obey perfect scale invariance. Given a point and a scale selection operator we compute the operator responses for a set of scales of σ_n and select the local extremum of the responses. Laplacian-of-Gaussian finds the highest percentage of correct characteristic scales by this equation:

$|LoG(x, \sigma_n)| = \sigma_n^2 |L_{xx}(x, \sigma_n) + L_{yy}(x, \sigma_n)|$

LoG is good at detecting blob and provides a good estimation of characteristic scale for other structures such as corners, edges, ridges and multi-junctions.

7.3. Harris Laplacian Detector

Harris Laplacian Detector could detect many representative key points on images. It uses the scale-adapted Harris function to localize points in scale-space and then selects the points for which Laplacian of Gaussian attains a maximum over scale. We detect the initial points with the multi-scale Harris detector and build the scale-space representation with the Harris function then detect local maxima at each scale level. After that we verify LoG on initial points whether it attains a maximum at the scales of points and ranged in specified area. In this way, we obtain a set of characteristic points with associated scales. Harris Laplacian Detector provides a compact and representative set of points which are characteristic in the image and in the scale dimension.

7.4. RIFT Descriptor on Key Points

RIFT is short for Rotation-Invariant Feature Transform. The patch besides the key point is divided into concentric rings of equal width and a gradient orientation is computed within each ring. Similar to HoG descriptor, we calculate histogram in these concentric rings coordinated by distance and angle. In order to maintain rotation invariance, orientation is measured at each point relative to the direction pointing outward from the center. In our framework, we use four rings and eight histogram orientations, getting 32-dimensional descriptors.

Figure show here to represent RIFT descriptor

7.5. K-means Clustering, Building Dictionary

We get a large size of key points descriptor using Harris Laplacian Detector and RIFT descriptor. In order to build a standard dictionary, we clusters these descriptors into K partitions by K-means clustering algorithm. K-means clustering is to cluster n key points descriptors into K (K = 2000 in our framework) partitions. Its purpose is to minimize this function

$V = \sum_{i=1}^{k} \sum_{x_j \in S_j} (x_j - \mu_i)^2$

where there are K clusters S_i and μ_i is the centroid of all the points x_j.

Initially, we randomly partition n key points descriptors into K partitions and get K initial centroids. After that we construct a new partition by associating each point with the closest centroid. Then the new centroids are recalculated by new clusters and we use this iterative algorithm to get final converged centroids. We build K length descriptor dictionary by this process.

7.6. Data Set Representation on Dictionary

We extract corresponding Hair, Skin and Face data set from our classification images. From these data set we extract key points and get descriptors by the same method above. Then we transform these features by dictionary mapping into same length size votes vectors. We input these vectors into SVMlight and get weak classifiers for Hair, Skin and Face.

By these steps, we use x training positive set and y training negative set to build hair classifier getting accuracy n%; x training positive set and y training negative set to build skin classifier getting accuracy n%; x training positive set and y training negative set to build face classifier getting accuracy n%. We draw a table to show the comparison between these three weak classifiers.

Figure shows the comparison here.

These weak classifiers are useful to get combined strong classifiers.

8. Torso and Leg Weak Classifier

Torso and Leg are parts of human and have nearly same features as human body. So as to Torso and Leg classifier, we use the same pipeline with human classifier but using different parameters which will be specifically discussed later. In human classifier we choose cell = 100 and block = 2*2, and now we change the parameter into cell = 50 and block = 2*2. In the specified area, Torso and Leg could attain a good accuracy on testing set, x% and y% respectively. Since Torso and Leg surely appear in human images, they are more convincing than Hair, Skin and Face.

Some of Torso and Leg sample images are shown here:
Figure to show sample images.

9. Combination of Different Classifiers

Now we have built multi weak classifiers on the selected set. Our next task is to combine these different classifiers using a linear model and choose the parameters to get the best accuracy x% on the training set. By this model and multi classifiers results we attain the final accuracy x% on testing set. Our linear model is as follows:

Figure shows this model here but needs more considerations.

10. Human Classification

Based on such classifiers and steps we build a strong classifier to get a high accuracy x% on testing set 2000*1200. This is like a cascade classifier to improve the accuracy. Here we draw a table to compare the performance before cascade and after cascade.

Figure here shows comparison.

11. Human Detection

We are now able to detect people in standard 2000*1200 high resolution images. Our next step is to detect people in large images containing a lot of people with different sizes and scales. Here we use sliding window detection algorithms with different scale parameters to detect people in the images. The sample final detection result is shown here. More final detection results are discussed in experiment part.

12. Experiment Results and Conclusions

In this section we shows the experiment results as described above.

Figure 1 shows the relation between cell size, block size and final accuracy. From figure 1 we could see that Rough Human Classification gets the best accuracy when cell = 100 and block = 2*2. In fact the larger cell the more details will disappear leading to low accuracy. In the contrary, the smaller cell the less details will disappear leading to generally high accuracy. But when cell size is too small it is a time-consuming process to do rough human classification. So the best parameter should be cell = 100 and block = 2*2.

Figure 2 shows comparison between windows size, detection step and detection accuracy when detecting people in detection images. It is better to select parameters using window = x, step = y detecting human in detection images.

Figure 3 shows the relation between different thresholds and final accuracy. It is shown that when threshold equals x the final accuracy achieves best. The threshold specifies the selected set to be further processed by multi classifiers. The higher threshold selects more data set which require more time while the lower threshold selects less data set requiring

less time. Actually, Hair, Skin, Face, Torso and Leg classifiers do not have distinctively higher accuracy than Rough Human Classifier. So the threshold should not be too large.

Figure 4 to Figure 7 shows the relation between five weak classifiers prediction and actual value. The sequence is Hair, Skin, Face, Torso and Leg. We could see that Hair and Skin are not so consistent; Face is a little better and Torso and Leg are much better to give correct predictions. The reasons are easy to find because we are not sure to find Hair, Skin and Face in the head area while we are sure Torso and Leg would appear in the human images. So Torso and Leg prediction results are more consistent compared with actual values.

Figure 5 shows different parameters combination results by multi classifiers versus final accuracy. The parameters are selected by repetitive trying. It is shown that x, y, z are the best parameters and get the best accuracy results. In fact linear model is the simplest model and a better complicated model is necessary to ensure the final improvement.

Figure 6 shows performance comparisons between different key points detectors: SIFT, Harris, Harris Laplace. Harris Laplace is relatively a better method to detect key points in our framework.

Figure 7 shows performance comparisons between different key points descriptors: SIFT-Descriptor, HoG-Descriptor, Spin-Descriptor, RIFT-descriptor. We could see RIFT-Descriptor is relatively a better descriptor to represent key points for classification.

Figure 8 shows comparison between different cell size and block size when classifying Torso and Leg. From the figure cell = 50 and block = 2*2 are the best parameters to build a good Torso and Leg classifier.

More figures and discussions will be added here.

13. Demo System

Our demo system is to integrate all algorithms and show results by MATLAB GUI. There are some figures to show the running process: From figure 9 - 14, it is our final demo system implemented by MATLAB GUI to show the excellent performance of our algorithm. Figure 9 is the interface with menu items and showing area. Figure 10 is loading training data and testing data to show in the axis objects. Figure 11 is the rough classification results using rough human classifier. Figure 12 is the selected area including head area, torso area and leg area. Figure 13 shows combination results based on multi weak classifier results. Figure 14 shows the detection results using sliding windows detection framework. More figures will be added here.

14. Conclusion and Future Work

In our framework, First we build a totally new data set by ourselves. All the images are photographed at Stanford and

reserved for research purpose only. Second we classify our data set using human classifier. Third we select undecided data set for further processing. Forth we build new weak classifiers on selected set. Forth we combine these results to get a better accuracy. Finally we use similar framework, varied parameters and varied sliding window algorithm to detect people in detection images with multi people. We have finished a whole system to classify human and detect human in high resolution images. Our current algorithm uses cascade classifiers to finalize the classification results. We can get better final results since there are still many improvements needed.

For future work. we could add more standard data set into our training set and testing set to improve robustness. During the threshold selection, we could further incorporate a better selection mechanism to get the best threshold between efficiency and accuracy. After multi classifiers results a better combination mechanism is needed to fully utilize classifier results. In the sliding window algorithm we could get a faster and more effective method. We could build a stronger Hair, Skin, Face, Torso and Leg classifiers using a novel feature, method and framework. In other words it is possible to use more features to get a better classification results especially for hair and skin. Hair and skin are hard to get perfect results on testing images even though it gets a high accuracy on training set and testing set.

15. 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.

15.1. 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.

15.2. 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 1-3/8 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 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 1 and 2. Short captions should be centred. 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.

15.3. Footnotes

Please use footnotes[1] 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.

15.4. 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 [4]. Where appropriate, include the name(s) of editors of referenced books.

---

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

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

Table 1. Results. Ours is better.

15.5. 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. 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}
```

15.6. 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.

16. 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.

Please direct any questions to the production editor in charge of these proceedings at the IEEE Computer Society Press: Phone (714) 821-8380, or Fax (714) 761-1784.

References

[1] A. Alpher, , and J. P. N. Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003. 3

[2] A. Alpher, , J. P. N. Fotheringham-Smythe, and G. Gamow. Can a machine frobnicate? *Journal of Foo*, 14(1):234–778, 2004. 3

[3] A. Alpher. Frobnication. *Journal of Foo*, 12(1):234–778, 2002. 3

[4] Authors. The frobnicatable foo filter, 2006. ECCV06 submission ID 324. Supplied as additional material eccv06.pdf. 2, 3, 7

[5] Authors. Frobnication tutorial, 2006. Supplied as additional material tr.pdf. 2