Part Association for Pedestrian Detection in Crowd Scene

Anonymous ECCV submission

Paper ID ***

Abstract. We propose the novel pedestrian detection method for crowded scenes which can provide an exact visible appearance of partially occluded pedestrians. Although the state-of-the art detectors based on a single pedestrian have achieved a great performance on moderate scenes, they have failed on severely occluded pedestrians. Thus, a numbers of alternative methods for crowded scene have been proposed, but many of them just modeled the spatial relationship between bounding boxes of each pedestrian to reduce the number of missed pedestrian. Furthermore, they require the additional post processing to remove redundant detection results. In this paper, we propose the pedestrian detection method associating visible parts according to each pedestrian. We formulate this association problem as a Binary Quadratic Programming (BQP) which is able to solve the occlusion reasoning and the detection redundancy problem at once. Experiments on the public benchmark validate the robust performance of the proposed method.

Keywords: pedestrian detection, occlusion reasoning, crowded scene analysis

1 Introduction

The pedestrain detection is essential for many practical applciations like the self-driving car, the visual analytics, robotics, the visual survaillance and etc. Futhermore, the performance of those applications are deeply affected by the quality of pedestrian detections. Thus, the pedestrian detection have been studied widely for the last decade. Many of recently proposed methods achieved a good performance on the moderate scene [**?**,**?**,**?**,**?**]. But they have failed on crowd scenes because they only modeled an isolated pedestrian so they cannot handle the significant variation at the appearance of pedestrians caused by occlusions between them. Moreover, they have lost true positive detections during the post processing for removing redundant detecions like Non-Maximum Suppression (NMS).

In recent years, methods model the partial occlusion of single pedestrian are proposed [**?**, **?**]. They learned several models represent various partial occlusion of a pedestrian and applied the models to target image altogether. This kind of method can detect more pedestrians than the previous methods. But they have difficulty to seperate the false positive detections from the detections of real

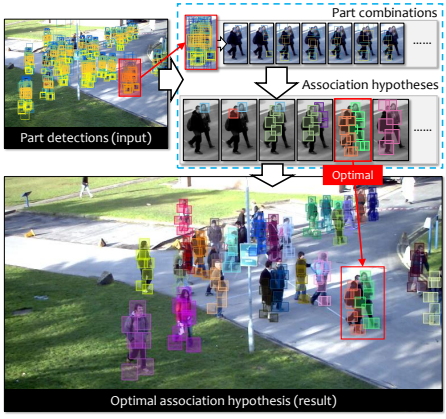

Fig. 1. The overall scheme of the proposed method.

pedestrians because they do not consider the occlusion relationships between pedestrians.

The alternative methods modeled the mutual occlusions of pedestrians [**?**, **?**]. They modeled typical cases of mutual occlusion scenarios between pedestrians by multiple classifiers. They achieved a good performance on several scenes which contain the occlusion patterns the classifiers learned. But there are the infinite number of possible occlusion scenarios, so modeling all the occlusions with several classifiers is impractical. Furthermore, they cannot handle more complex occlusions related with more than three pedestrians on severly crowd scenes.

Yan *etal*. [**?**] proposed the method based on the optimization for high density crowd scene. In [**?**] the spatial relationships between bounding boxes of mutually occluded pedestrians are modeled. The modeling information was used in the calculation of scores in the optimization step. But their detections do not contain any information about the appearance of each pedestrian because their method does not have explicit occlusion reasoning. Moreover, the learned model was the crucial prior information about the distribution of bounding boxes since they used scenes in the training which were captured by exactly same camera with the test set. This kind of the strong prior informations are unrealistic in general situations.

In this paper, we propose the novel pedestrian detection method which considers the overall occlusions in the target image at once. In our method, we explicitely infer the body part occlusions to distinguish detections of partially occluded pedestrians in false positive detections and also to provide the confidential appearance of each pedestrian. For the reasoning of body part occlusions, we formulate the pedestrian detection problem as the part association problem and solve the problem with the Binary Quadratic Programming (BQP). The proposed BQP not only does the occlusion reasoning, but also resolves the problem of redundant detections. In addition, the prior information about occluding area in the target image can be easily unified to our BQP when it is available.

Experiments on the public benchmarks show that the proposed method has the robust detection performance and capability to provide the exact appearances of each pedestrian.

2 Paper formatting

2.1 Language

All manuscripts must be in English.

2.2 Paper length

Papers submitted for review should be complete. The length should match that intended for final publication. Papers accepted for the conference will be allocated 14 pages (plus references) in the proceedings. Note that the allocated 14 pages do not include the references. The reason for this policy is that we do not want authors to omit references for sake of space limitations.

Papers with more than 14 pages (excluding references) will be rejected without review. This includes papers where the margins and formatting are deemed to have been significantly altered from those laid down by this style guide. 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 14 pages if it is reviewed in 16.

2.3 Paper ID

It is imperative that the paper ID is mentioned on each page of the manuscript. The paper ID is a number automatically assigned to your submission when registering your paper submission on CMT.

2.4 Line numbering

All lines should be numbered, as in this example document. This makes reviewing more efficient, because reviewers can refer to a line on a page. If you are preparing a document using a non-LaTeX document preparation system, please arrange for an equivalent line numbering.

2.5 Mathematics

Please number all of your sections and displayed equations. Again, this makes reviewing more efficient, because reviewers can refer to a line on a page. Also, 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 line numbering will not be present in the final copy, so is not an alternative to equation numbers). Some authors might benefit from reading Mermin's description of how to write mathematic.

3 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 technical reports).

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 [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 [1] as additional material and cite it as

> 1. Authors. "The frobnicatable foo filter", BMVC 2014 Submission ID 324, Supplied as additional material `bmvc14.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 [2]". 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 ECCV 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.

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, 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.

For sake of anonymity, it's recommended to omit acknowledgements in your review copy. They can be added later when you prepare the final copy.

4 Manuscript Preparation

This is an edited version of Springer LNCS instructions adapted for ECCV 2016 first paper submission. You are strongly encouraged to use LaTeX2$_\varepsilon$ for the preparation of your camera-ready manuscript together with the corresponding Springer class file `llncs.cls`.

We would like to stress that the class/style files and the template should not be manipulated and that the guidelines regarding font sizes and format should be adhered to. This is to ensure that the end product is as homogeneous as possible.

4.1 Printing Area

The printing area is 122 mm × 193 mm. The text should be justified to occupy the full line width, so that the right margin is not ragged, with words hyphenated as appropriate. Please fill pages so that the length of the text is no less than 180 mm.

4.2 Layout, Typeface, Font Sizes, and Numbering

Use 10-point type for the name(s) of the author(s) and 9-point type for the address(es) and the abstract. For the main text, please use 10-point type and single-line spacing. We recommend using Computer Modern Roman (CM) fonts, Times, or one of the similar typefaces widely used in photo-typesetting. (In these typefaces the letters have serifs, i.e., short endstrokes at the head and the foot of letters.) Italic type may be used to emphasize words in running text. Bold type and underlining should be avoided. With these sizes, the interline distance should be set so that some 45 lines occur on a full-text page.

Headings. Headings should be capitalized (i.e., nouns, verbs, and all other words except articles, prepositions, and conjunctions should be set with an initial capital) and should, with the exception of the title, be aligned to the left. Words joined by a hyphen are subject to a special rule. If the first word can stand alone, the second word should be capitalized. The font sizes are given in Table 1.

Table 1. Font sizes of headings. Table captions should always be positioned *above* the tables. The final sentence of a table caption should end without a full stop

Heading level	Example	Font size and style
Title (centered)	**Lecture Notes ...**	14 point, bold
1st-level heading	**1 Introduction**	12 point, bold
2nd-level heading	**2.1 Printing Area**	10 point, bold
3rd-level heading	**Headings.** Text follows ...	10 point, bold
4th-level heading	*Remark.* Text follows ...	10 point, italic

Here are some examples of headings: “Criteria to Disprove Context-Freeness of Collage Languages”, “On Correcting the Intrusion of Tracing Non-deterministic Programs by Software”, “A User-Friendly and Extendable Data Distribution System”, “Multi-flip Networks: Parallelizing GenSAT”, “Self-determinations of Man”.

Lemmas, Propositions, and Theorems. The numbers accorded to lemmas, propositions, and theorems etc. should appear in consecutive order, starting with the number 1, and not, for example, with the number 11.

4.3 Figures and Photographs

Please produce your figures electronically and integrate them into your text file. For LaTeX users we recommend using package `graphicx` or the style files `psfig` or `epsf`.

Check that in line drawings, lines are not interrupted and have constant width. Grids and details within the figures must be clearly readable and may not be written one on top of the other. Line drawings should have a resolution of at least 800 dpi (preferably 1200 dpi). For digital halftones 300 dpi is usually sufficient. The lettering in figures should have a height of 2 mm (10-point type). Figures should be scaled up or down accordingly. Please do not use any absolute coordinates in figures.

Figures should be numbered and should have a caption which should always be positioned *under* the figures, in contrast to the caption belonging to a table, which should always appear *above* the table. Please center the captions between the margins and set them in 9-point type (Fig. 2 shows an example). The distance between text and figure should be about 8 mm, the distance between figure and caption about 5 mm.

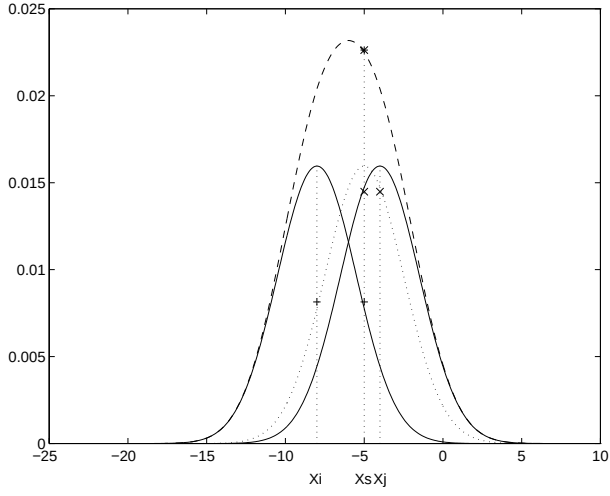

Fig. 2. One kernel at x_s (*dotted kernel*) or two kernels at x_i and x_j (*left and right*) lead to the same summed estimate at x_s. This shows a figure consisting of different types of lines. Elements of the figure described in the caption should be set in italics, in parentheses, as shown in this sample caption. The last sentence of a figure caption should generally end without a full stop

If possible (e.g. if you use LaTeX) please define figures as floating objects. LaTeX users, please avoid using the location parameter "h" for "here". If you have to insert a pagebreak before a figure, please ensure that the previous page is completely filled.

4.4 Formulas

Displayed equations or formulas are centered and set on a separate line (with an extra line or halfline space above and below). Displayed expressions should be numbered for reference. The numbers should be consecutive within the contribution, with numbers enclosed in parentheses and set on the right margin. For example,

$$\psi(u) = \int_0^T \left[\frac{1}{2}\left(\Lambda_0^{-1}u, u\right) + N^*(-u)\right] dt \tag{1}$$

$$= 0? \tag{2}$$

Please punctuate a displayed equation in the same way as ordinary text but with a small space before the end punctuation.

4.5 Footnotes

The superscript numeral used to refer to a footnote appears in the text either directly after the word to be discussed or, in relation to a phrase or a sentence, following the punctuation sign (comma, semicolon, or full stop). Footnotes should appear at the bottom of the normal text area, with a line of about 2 cm in TEX and about 5 cm in Word set immediately above them.[1]

4.6 Program Code

Program listings or program commands in the text are normally set in typewriter font, e.g., CMTT10 or Courier.
Example of a Computer Program

```
program Inflation (Output)
  {Assuming annual inflation rates of 7%, 8%, and 10%,...
   years};
   const
     MaxYears = 10;
   var
     Year: 0..MaxYears;
     Factor1, Factor2, Factor3: Real;
   begin
     Year := 0;
     Factor1 := 1.0; Factor2 := 1.0; Factor3 := 1.0;
     WriteLn('Year  7% 8% 10%'); WriteLn;
     repeat
       Year := Year + 1;
```

[1] The footnote numeral is set flush left and the text follows with the usual word spacing. Second and subsequent lines are indented. Footnotes should end with a full stop.

```
        Factor1 := Factor1 * 1.07;
        Factor2 := Factor2 * 1.08;
        Factor3 := Factor3 * 1.10;
        WriteLn(Year:5,Factor1:7:3,Factor2:7:3,Factor3:7:3)
      until Year = MaxYears
end.
```

(Example from Jensen K., Wirth N. (1991) Pascal user manual and report. Springer, New York)

4.7 Citations

The list of references is headed "References" and is not assigned a number in the decimal system of headings. The list should be set in small print and placed at the end of your contribution, in front of the appendix, if one exists. Please do not insert a pagebreak before the list of references if the page is not completely filled. An example is given at the end of this information sheet. For citations in the text please use square brackets and consecutive numbers: [3], [4], [5] ...

5 Conclusions

The paper ends with a conclusion.

Page 10 of the manuscript.

Page 11 of the manuscript.

Page 12 of the manuscript.

Page 13 of the manuscript.

Page 14 of the manuscript.

Page 15 of the manuscript.

Page 16 of the manuscript. This is the last page of the manuscript.

Now we have reached the maximum size of the ECCV 2016 submission (excluding references). References should start immediately after the main text, but can continue on p.15 if needed.

References

1. Authors: The frobnicatable foo filter (2014) BMVC14 submission ID 324. Supplied as additional material `bmvc14.pdf`.
2. Authors: Frobnication tutorial (2014) Supplied as additional material `tr.pdf`.
3. Alpher, A.: Frobnication. Journal of Foo **12**(1) (2002) 234–778
4. Alpher, A., , Fotheringham-Smythe, J.P.N.: Frobnication revisited. Journal of Foo **13**(1) (2003) 234–778
5. Alpher, A., , Fotheringham-Smythe, J.P.N., Gamow, G.: Can a machine frobnicate? Journal of Foo **14**(1) (2004) 234–778