CVPR 2014
IEEE Conference on Computer Vision and Pattern Recognition 2014
June 17-19, 2014, Columbus, Ohio, USA

Reviews For Paper

Paper ID	1835
Title	Efficient feature extraction, encoding and classification\\ for action recognition

Masked Reviewer ID: Assigned_Reviewer_10

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	This paper provides an efficient way to extract local video features using the motion information from MPEG video compression. They further explore a few feature encoding methods (histogram, Fisher vectors and VLAD) in their experiments. They also show that their approach improves the speed of video recognition pipeline for one to two orders of magnitude with small performance drops.
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think	The main contribution of the paper is that it shows using the less accurate motion information that exists in MPEG encoding is not that bad for action recognition and achieves good accuracy, but at the same time is orders of magnitude faster. I think if they provide their code, it will be useful for the community.

this is the case.	
Paper Weaknesses. Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to CVPR and that a theoretical paper need not have experiments. It is not okay to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not reasonable to ask for comparisons with unpublished papers and papers published after the CVPR deadline.	Use of video compression features to approximate motion flows is interesting. However, I think the paper is not novel from the approach point of view. Apart from the proposed motion descriptors from video compression, the rest of the paper looks like a survey comparing several existing flow estimation methods, feature encoding methods, etc.
Preliminary Rating. Please rate the paper according to the	

following choices. Oral: these are papers whose quality is in the top 10% of the papers at CVPR. Examples include a theoretical breakthrough with no experiments; an interesting solution to a new problem; a novel solution to an existing problem with solid experiments; or an incremental paper that leads to dramatic improvements in performance. Oral/Poster: these are very strong papers, which may have one weakness that makes you unsure as to whether they should be oral or poster. Poster: these are strong papers, which have more than one weakness. For example, a well-written paper with solid experiments, but incremental; a paper on a well studied problem with solid theory, but weak experiments; or a novel paper with good experiments, but poorly written. Weak Reject: these are papers that have some promise, but they would be better off by being revised and resubmitted. Strong Reject: these are papers	Poster	

that have major flaws, or have been done before.	
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making. There is no need to summarize the paper or reviews. If you have additional concerns that were not included in the reviews, please be sure and include them as well.	Obviously the methodology is not novel, but the reason I am voting for a poster is that their code might be useful for people who want to do large scale action recognition and need a very fast method.
Confidence. Write "Very Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your	Confident.

conclusions. In the latter case, please provide details.	

Masked Reviewer ID: Assigned_Reviewer_2

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	This paper proposed an efficient video feature extraction method for action recognition. As the optical flow estimation is a time-consuming process which limits the speed of feature extraction, the motion information provided by the video compression is used in the proposed approach. In addition, this paper evaluated both Fisher Vector and VLAD for action recognition.
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	The proposed approach can improve the efficiency of action recognition because of the speed-up in feature extraction, without sacrificing too much on performance. Therefore, it may be useful for action recognition when employing the compressed videos in MPEG format.
Paper Weaknesses. Please discuss	

the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to CVPR and that a theoretical paper need not have experiments. It is not okay to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not reasonable to ask for comparisons with unpublished papers and papers published after the CVPR deadline.	1. The feature extraction is speeded-up as the MPEG motion is used to replace the optical flow. However, it is unclear whether the propose method can work with videos that are not MPEG format. 2. The major speed gain comes from the motion information in the compressed video. However, whether such a replacement of optical flow is suitable or not should be justified through more experiments. The paper gives some interesting illustrations in Figure 1 and 2. However it may not be convincing enough, as there are different types of MPEG motion estimation algorithms and different objectives. Also for optical flow estimation, there are also fast methods that can provide coarse optical flow map and the method should consider comparing with. 3. The evaluation of Fisher Vector and VLAD for action recognition is interesting but cannot be considered as the contribution of the paper. Mihir Jain et.al. have already reported the performance of VLAD for action recognition in “Better exploiting motion for better action recognition” CVPR 2013. 4. The evaluation in Table 1 shown that there is a 5% (in average) gap between the proposed feature and the dense trajectory feature. Considering this significant performance degradation, it is necessary to give detailed analysis of the trade-off between accuracy and efficiency. For example, by adjusting compression algorithms and parameters, the performance of the proposed feature may be improved, even with higher computational cost.
Preliminary Rating. Please rate the paper according to the following choices. Oral: these are papers whose quality is in the top 10% of the	

papers at CVPR. Examples include a theoretical breakthrough with no experiments; an interesting solution to a new problem; a novel solution to an existing problem with solid experiments; or an incremental paper that leads to dramatic improvements in performance. Oral/Poster: these are very strong papers, which may have one weakness that makes you unsure as to whether they should be oral or poster. Poster: these are strong papers, which have more than one weakness. For example, a well-written paper with solid experiments, but incremental; a paper on a well studied problem with solid theory, but weak experiments; or a novel paper with good experiments, but poorly written. Weak Reject: these are papers that have some promise, but they would be better off by being revised and resubmitted. Strong Reject: these are papers that have major flaws, or have been done before.	Weak Reject

Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making. There is no need to summarize the paper or reviews. If you have additional concerns that were not included in the reviews, please be sure and include them as well.	Although the use of MPEG motion flow for video feature extraction is interesting, both the novelty and relevance of the technical contributions of this paper are weak for the CVPR conference.
Confidence. Write "Very Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your conclusions. In the latter case, please provide details.	very confident

Masked Reviewer ID: Assigned_Reviewer_9

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	This paper explores features from motion flow and compress field for efficient action recognition.
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	The speed is very good.
Paper Weaknesses. Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient	

experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to CVPR and that a theoretical paper need not have experiments. It is not okay to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not reasonable to ask for comparisons with unpublished papers and papers published after the CVPR deadline.	The writing could be further improved.
Preliminary Rating. Please rate the paper according to the following choices. Oral: these are papers whose quality is in the top 10% of the papers at CVPR. Examples include a theoretical breakthrough with no experiments; an interesting	

solution to a new problem; a novel solution to an existing problem with solid experiments; or an incremental paper that leads to dramatic improvements in performance. Oral/Poster: these are very strong papers, which may have one weakness that makes you unsure as to whether they should be oral or poster. Poster: these are strong papers, which have more than one weakness. For example, a well-written paper with solid experiments, but incremental; a paper on a well studied problem with solid theory, but weak experiments; or a novel paper with good experiments, but poorly written. Weak Reject: these are papers that have some promise, but they would be better off by being revised and resubmitted. Strong Reject: these are papers that have major flaws, or have been done before.	Poster
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current	I am very impressed by the fast speed of the proposed method. As shown in Tables 1, 3, 4, the speed of the proposed method is about 100 times faster than then dense trajectory method [36] while the performance is comparable. However, I feel this paper could still be improved in the following aspects:

opinion on the paper. Please summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making. There is no need to summarize the paper or reviews. If you have additional concerns that were not included in the reviews, please be sure and include them as well.	(1) Compare with the speed of other methods. It is well known that dense trajectory is among the slowest action recognition models. I am interested to know whether the proposed method is faster than other faster methods, e.g., Hu et al, Action detection in complex scenes with spatial and temporal ambiguities, ICCV'09 Shi et al, Sampling strategies for real-time action recognition, CVPR'13 (2) Most of the experiments are compared with [Wang et al, IJCV'13]. It is worth mentioning that the same group has published a ICCV paper with improved performance: H. Wang and C. Schmid. Action Recognition with Improved Trajectories. ICCV, 2013 (3) Some paragraphs could be polished or improved. For example, it seems to me sections 4.1, 4.2, and 4.3 are mostly previous works, and it is not clear to me where the novel contribution is.
Confidence. Write “Very Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), “Confident” to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and “Not Confident” to stress that that you feel some doubt about your conclusions. In the latter case, please provide details.	confident