

From: ls@cs.brown.edu
To: [Stefan Roth](#); [Michael Isard](#); [Hulya Yalcin](#); [Michael J.Black](#)
Cc: black@opticalflow.com
Subject: Fwd: Reviews for Paper #1109
Date: Friday, February 13, 2004 3:04:10 PM

----- Forwarded message from reynolds@cc.gatech.edu -----

Date: Fri, 13 Feb 2004 19:37:42 GMT
From: reynolds@cc.gatech.edu
Reply-To: reynolds@cc.gatech.edu
Subject: Reviews for Paper #1109
To: ls@cs.brown.edu, cvprcl@cc.gatech.edu

Dear Leonid,

In preparation for the author response period, please find below the reviews we have received for the following paper:

Paper ID: 1109
Title: Tracking Loose-limbed People

These reviews were written during the initial reviewing phase and then discussed by the Program Committee Members and Area Chair(s) in charge of your paper.

Please note that the set of attached reviews is not final. While the content of the reviews below will not change, your paper will receive a summary written by an Area Chair and may receive additional reviews before final decisions are made. Papers for which only two reviews are currently available will receive at least one more review.

Instructions for the Author Response Period

1. The author response period will begin on Feb 12, 2004 and will end on Feb 17th at 10pm EST.
2. Responses are limited to 4000 characters of plain text.
3. Author responses must be entered directly on the website.
Go to the 'My Papers' page after logging in and click on the appropriate link for each of your papers.
4. Your response will be read only by the Area Chair(s) in charge of your paper and by the Program Chairs.
Your response will NOT be communicated verbatim to the reviewers of your paper, but may result in further discussion between the Area Chair and those reviewers.

IMPORTANT:

- AUTHOR RESPONSES ARE ENTIRELY VOLUNTARY. FAILURE TO SUBMIT A RESPONSE WILL NOT BE CONSIDERED NEGATIVELY DURING FINAL DECISION MAKING

- DO NOT INCLUDE ANY INFORMATION IN YOUR RESPONSE THAT MAY REVEAL AUTHOR IDENTITIES!!! The entire decision process will be conducted double-blind. IF THE RESPONSE CONTAINS ANY IDENTIFYING INFORMATION (WEBSITES, EMAILS, ETC) IT WILL BE DELETED IN ITS ENTIRETY.

- DO NOT EMAIL your response to the Program Chairs; any response sent via email will be discarded without review.

* AUTHOR RESPONSE PERIOD ENDS FEB 17 AT 10PM EST *

* *

* THIS IS A HARD DEADLINE---NEITHER THE WEBSITE *

* NOR THE PROGRAM CHAIRS WILL ACCEPT LATE RESPONSES *

We will be in touch again after a final decision is reached on your paper.

Regards

Aaron Bobick
Rama Chellappa
CVPR 2004 Program Chairs

REVIEW #1

** SUMMARY AND CONTRIBUTIONS:

This paper describes a method for tracking people using multiple cameras. A method for initializing the tracker is incorporated into the framework.

Tracking is posed as an inference problem in a graphical model with loops, containing spatial and temporal links. Potentials between limbs are modeled as mixtures of Gaussians. Likelihoods for individual limbs are represented by Gibbs models.

Parameters for both are learned from a training set. Inference is performed using the PAMPAS algorithm for non-parametric belief propagation.

The weakest point of the paper is the experimental results. Results on tracking (with the aid of background subtraction) one sequence of a person walking slowly are presented. This experiment is not convincing in terms of the applicability of these methods. How useful would the learned distributions be if they were learned from more than just walking data? In particular, the lack of difficulty in handling the distributions of quaternions is suspicious.

The use of BP and particle filtering to perform tracking is a solid contribution. Although the results are lacking, using motion capture to present quantitative results is refreshingly scientific.

More details on what the training sets were for the limb detectors (how many people, frames, etc.) are needed.

It would be interesting to see this framework applied to do more detailed reasoning about occlusion.

Tracking and initialization are somewhat separate. What is the relationship between the part detectors and the likelihood $p(f|X_i)$?

**** NOVELTY:** [Somewhat original]

(applies PAMPAS, other established learning techniques to tracking people)

**** RELEVANCE:** [Of broad interest]

**** CLARITY OF PRESENTATION:** [Is clear enough]

Notation in equation (2) is confusing. Is $M(X_{ij})$ a transformation?

**** TECHNICAL CORRECTNESS:** [Probably correct (convinced but didn't check everything)]

**** REFERENCE TO PRIOR WORK:** [References good enough]

**** EXPERIMENTAL EVALUATION:** [Results unconvincing]

.....

**** OVERALL RECOMMENDATION:** [Probably Accept]

The use of BP and particle filtering to perform tracking is a solid contribution. Although the results are lacking, using motion capture to present quantitative results is refreshingly scientific.

.....

**** ADDITIONAL COMMENTS TO AUTHOR(S):**

REVIEW #2

**** SUMMARY AND CONTRIBUTIONS:**

A multiple-camera stationary-background person tracker is presented. A

graphical model for a "loose-limbed" person is learned by belief propagation. The spatial and temporal priors are gathered from data (using 3d mocap) and low-level part detectors ("shouters") are used as evidence. The results (for a single sequence) are compared with mocap ground truth.

**** NOVELTY:** [Somewhat original]

A nice probabilistic story, tying together several machine-learning techniques.

**** RELEVANCE:** [Of broad interest]

**** CLARITY OF PRESENTATION:** [Reads very well]

nice
but needs more figures of "shouters" and more results figures.

**** TECHNICAL CORRECTNESS:** [Probably correct (convinced but didn't check everything)]

**** REFERENCE TO PRIOR WORK:** [References good enough]

**** EXPERIMENTAL EVALUATION:** [Lacking in some respects]

This is very much a machine-learning paper: lots of nice, sound theory and then half a page on implementation and results plus a single example (although comparison with mocap ground truth is very nice!).

With such a general algorithm and such a rich setup -- 4 cameras, stationary background, training data -- the approach should be expected to generalize well across different people/clothes and different backgrounds. Results on a single sequence are very suspect (and I have a suspicion that the training for Sec 5 is done on frames from the very same sequence, which would be bad).

If the only problem is space constraints, it would be great if you could submit results for other actors and other backgrounds as part of the rebuttal -- it will greatly strengthen your paper!

.....

**** OVERALL RECOMMENDATION:** [Probably Accept]

The paper is well written and has a nice, coherent, probabilistic story. The experimental validation is quite poor but this is a very hard (and very interesting) area so even in this state the paper should probably be published. If the authors can submit more examples during rebuttal, I would raise my rating to "definitely accept".

.....

**** ADDITIONAL COMMENTS TO AUTHOR(S):**

I wonder how much do you rely on your shoulders to find your good parts initially? It seems that in the one example, the shoulders are doing very well, making your problem relatively easy. For example, the head is always found, which, of course, makes a lot of other things fall into place. Will the algorithm break if head is not found? How many limbs do you need to find? What's the breaking point?

REVIEW #3

**** SUMMARY AND CONTRIBUTIONS:**

A method for initializing and tracking 3D human body pose from multiple calibrated cameras. Uses learned (but apparently rather simplistic) body part detectors, a kinematic body model in which articulation is only enforced approximately, with particle filtering in a loopy Bayesian net to glue it all together.

**** NOVELTY:** [Somewhat original]

The various components have appeared before, but the combination is original and potentially useful.

**** RELEVANCE:** [Of sufficient interest]

**** CLARITY OF PRESENTATION:** [Is clear enough]

**** TECHNICAL CORRECTNESS:** [Probably correct (convinced but didn't check everything)]

**** REFERENCE TO PRIOR WORK:** [References good enough]

**** EXPERIMENTAL EVALUATION:** [Limited but convincing]

More detailed testing of the contributions of the various components would have been nice, but space obviously didn't allow this.

.....
**** OVERALL RECOMMENDATION:** [Probably Accept]

The method uses an interesting and topical combination of techniques,

and this is one of only a few 3D human trackers that can self-initialize. But the techniques have all appeared in some form before, and given that 4 calibrated cameras and background subtraction are used, I'm a little sceptical that all of this sophistication was really necessary.

.....

** ADDITIONAL COMMENTS TO AUTHOR(S):

----- End forwarded message -----

This message was sent using IMP, the Internet Messaging Program.