Forefront Object Segmentation for Moving Camera Sequences Based on Foreground-Background Probabilistic Models and Prior Probability Maps

Jaime Gallego, Pascal Bertolino
GIPSA-lab, Grenoble, France
Jaime.Gallego-Vila,pascal.bertolinog@gipsa-lab.grenoble-inp.fr

Abstract

• Monocular moving RGB camera where foreground object and background can present difficult situations like: camouflage, zoom effects, rotations, rigid/non-rigid shape.

• Use 3 region-based probabilistic models for foreground, near background surrounding the object and global background into MAP-MRF framework.

• Use prior probability maps obtained from the cumulative segmentation within temporal window.

1. Introduction

o Pixel-wise segmentation methods:
  • In general, not valid for moving camera sequences models difficult to build and update correctly

o Methods based on the evolution of different features of the image:
  • Estimate optical flow, motion saliency ...

o Our approach:
  Probabilistic framework

  ➢ Use region-based probabilistic models:
    • Foreground model gathers the information of the object.
    • Near Background models the information of the background that surrounds the object.
    • Global background describes more relevant color regions of each frame to avoid false detections due to background occlusions.

  ➢ Prior probability maps: cumulative knowledge of the object used to preserve its spatial shape.

  ➢ Use a Bayesian decision for pixel classification

  ➢ Apply regularization by using Graph cuts energy minimization

2. Probabilistic models

  • Combine foreground, near background and global background models into MAP-MRF framework:
    o Foreground and Near Background: Region based model Spatial Color Gaussian Mixture Model (SCGMM) in z=(R,G,B,X,Y) space.
    \[ P(z_i|l) = \sum_{k=1}^{K_l} \omega_k G_k(z_i, \mu_k, \Sigma_k) \]
    \[ l \in \{fg, bg\} \]

    o Global Background: Get the most representative regions of the background, and create a SCGMM with the R,G,B values.
    \[ P(z_i|\text{global bg}) = \sum_{l=1}^{N} \frac{Q_l}{Q} \omega_l G_{gb}(z_i, \mu_{gb}, \Sigma_{gb}) \]

3. Spatial Prior Probability Maps

  o Consecutive frames in a video sequence:
    • High degree of overlapping
    • Objects to segment present a moderate degree of change

  Take into account the history of the object segmentation into the classification process:

    • FIFO queue with the last J segmentation masks
    • Normalize the spatial domain of each mask by using the centroid position Correct overlapping of the J masks.

  o The spatial prior probability \( P_i(l) \) is formulated as:

    \[ P_i(l) = \frac{1}{J} \sum_{j=1}^{J} M_{i,l}(t-j), \quad l \in \{fg, bg\} \]

    \[ M_{i,l}(t-j) \in (1,0) \text{ for mask obtained for class } l \text{ in frame } (t-j) \]

4. MAP-MRF pixel classification

  • Bayesian pixel classification and graph cuts regularization: The likelihood for each pixel \( i \) is defined for each class as:

    \[ P(l_i|z_i) \propto P(z_i|l_i) P(l_i) \quad l \in \{fg, bg\} \]

    Where \( P(bg|z_i) = max(P(\text{near bg}|z_i), P(\text{global bg}|z_i)) \)

    • We consider a MRF framework: take into account neighborhood information. → Solved with Graph-cuts

5. Results & Conclusions

• Evaluation analyzing SegTrackv2 database.

Correct object segmentation reducing false positives, and false negatives detections also in those complicated scenes where camera motion, object changes and occlusions are present.

Sequence | Segmentation Technique | SPT+CSI | Key seg. | Bayes | Bayes p maps
--- | --- | --- | --- | --- | ---
Girl | 89.2 | 87.7 | 87.82 | 87.86 |
Birfall | 62.5 | 49.0 | 29.13 | 59.60 |
Cheetah-1 | 40.9 | 11.7 | 16.47 | 20.51 |
Cheetah-2 | 93.4 | 96.3 | 94.03 | 93.02 |
Parachute | 71.3 | 74.3 | 75.60 | 80.17 |
Monkeydog-1 | 18.9 | 4.9 | 48.02 | 48.29 |
Penguin-1 | 51.5 | 12.6 | 83.18 | 95.41 |
Penguin-2 | 76.5 | 11.3 | 80.35 | 89.35 |
Penguin-3 | 75.2 | 11.3 | 79.43 | 81.07 |
Penguin-4 | 57.8 | 7.7 | 73.80 | 80.62 |
Penguin-5 | 66.2 | 4.2 | 72.75 | 76.34 |

Download the software and see more results at
https://sites.google.com/site/jaimegallegovila/icip2014_bayesian_prior
http://www.gipsa-lab.grenoble-inp.fr/~pascal.bertolino/projects/readplay1/

*The authors would like to thank the French Région Rhônes-Alpes for its funding of this work in the context of the ReadPlay project.*