Kingston University London

PROFILE HIDDEN MARKOV MODELS FOR FOREGROUND OBJECT MODELLING

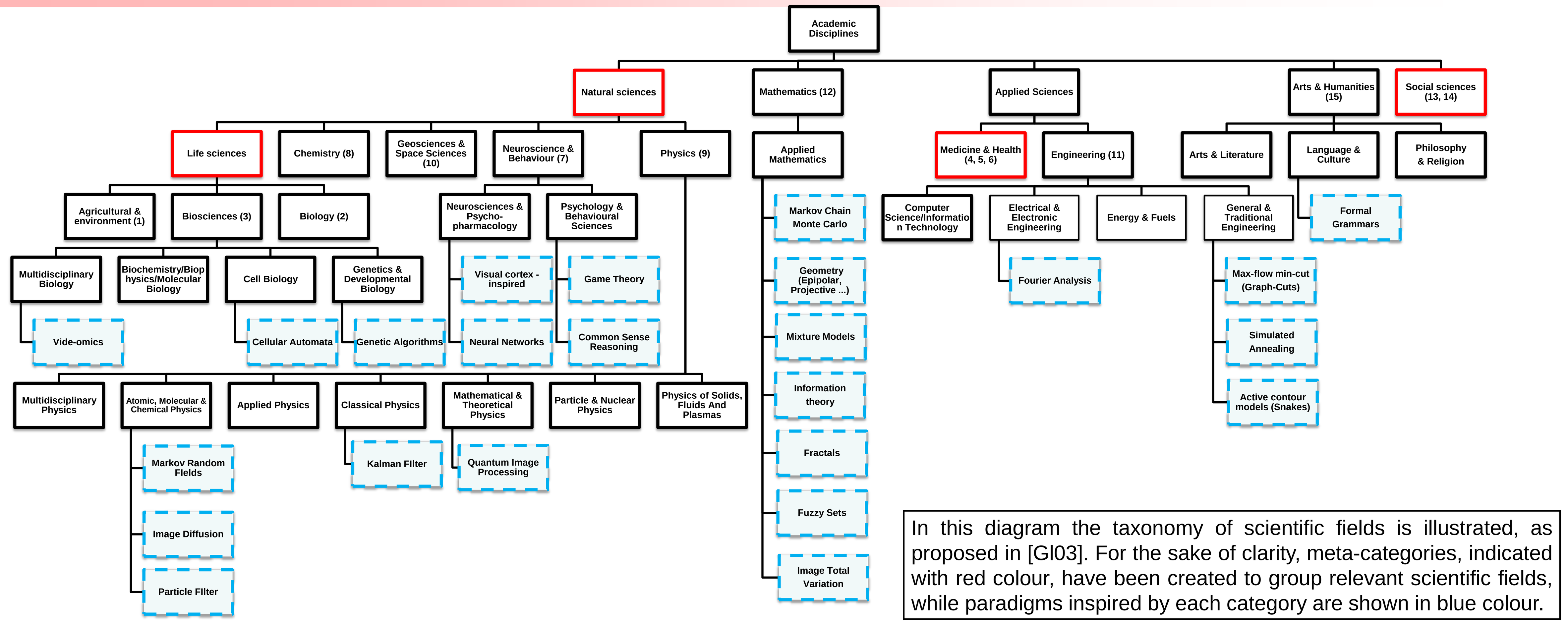
Ioannis Kazantzidis*, Francisco Florez-Revuelta , Jean-Christophe Nebel
*School of Computer Science and Mathematics, Kingston University, London
*ikazant@kingston.ac.uk

University of Alicante

Introduction

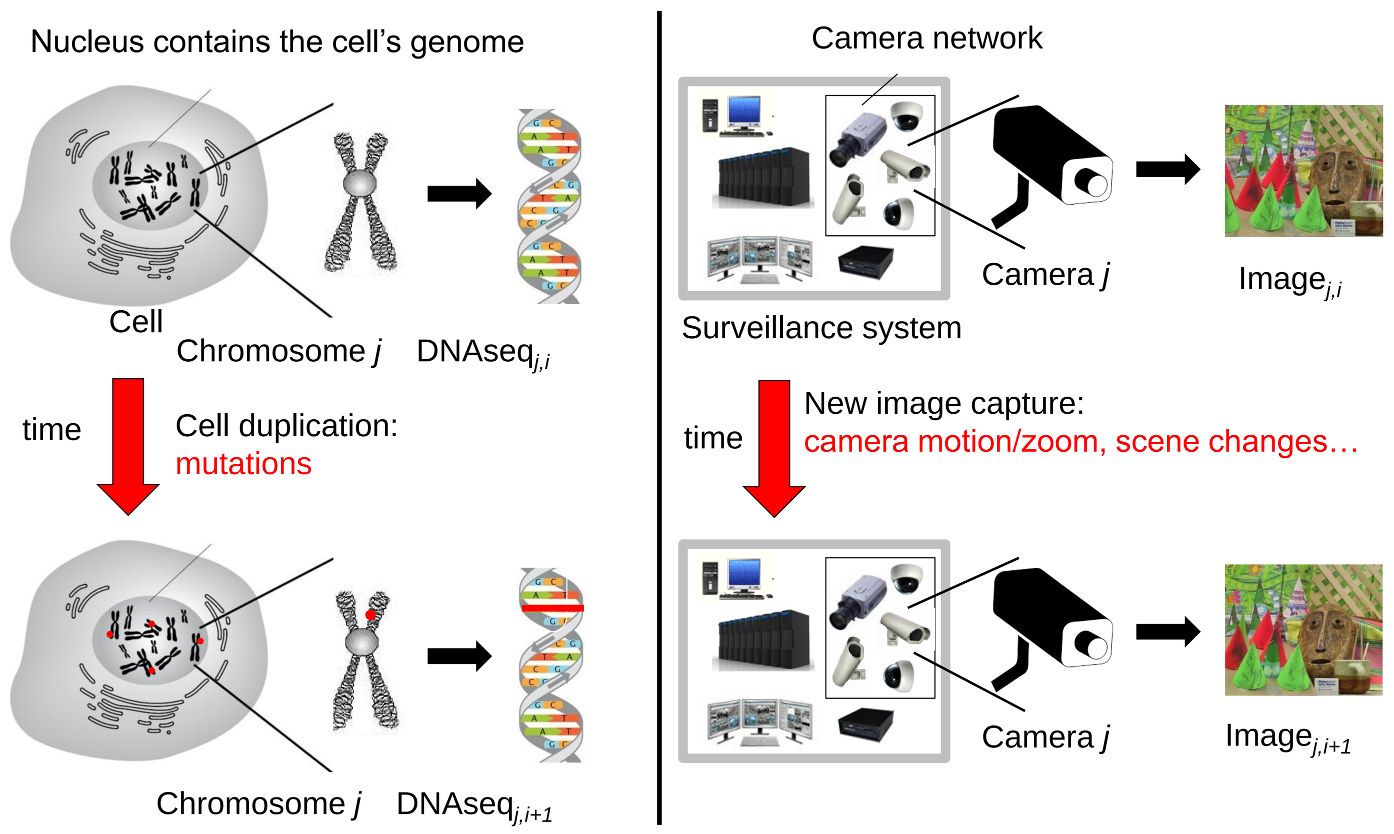
Challenges in video processing for visual surveillance such as shadows, reflections, illumination changes, occlusions, and weather conditions as well as moving objects and camera motion and vibration are common and induce variability for each pixel in an image. Current computer vision approaches rely on controlling the huge number of parameters involved in scene variability, by addressing those challenges by creating extra add-on modules in their processing pipeline. We believe a new strategy may be required underpinned by the realisation that scene variability should be the expected norm rather than an inconvenience to control.

Paradigms in Computer Vision



In this diagram the taxonomy of scientific fields is illustrated, as proposed in [Gl03]. For the sake of clarity, meta-categories, indicated with red colour, have been created to group relevant scientific fields, while paradigms inspired by each category are shown in blue colour.

Vide-omics: a genomics-inspired paradigm for video analysis



Analogy between (a) Cell duplication and (b) video capture by a surveillance system.

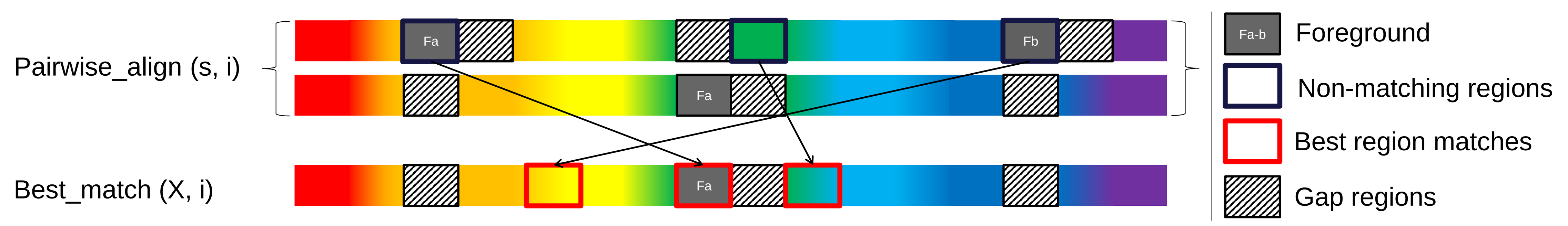
Substitution: Sensor noise, change in camera gain, change in scene illumination

Insertion: Change in camera angle and/or position revealing previously occluded data, more details in common field of view (zoom in), apparition of a new object in a camera's field of view after motion or zoom out

Deletion: Change in camera angle and/or position introducing new occlusions, less detail in common field of view (zoom out), disappearance of an object from a camera's field of view after motion or zoom in

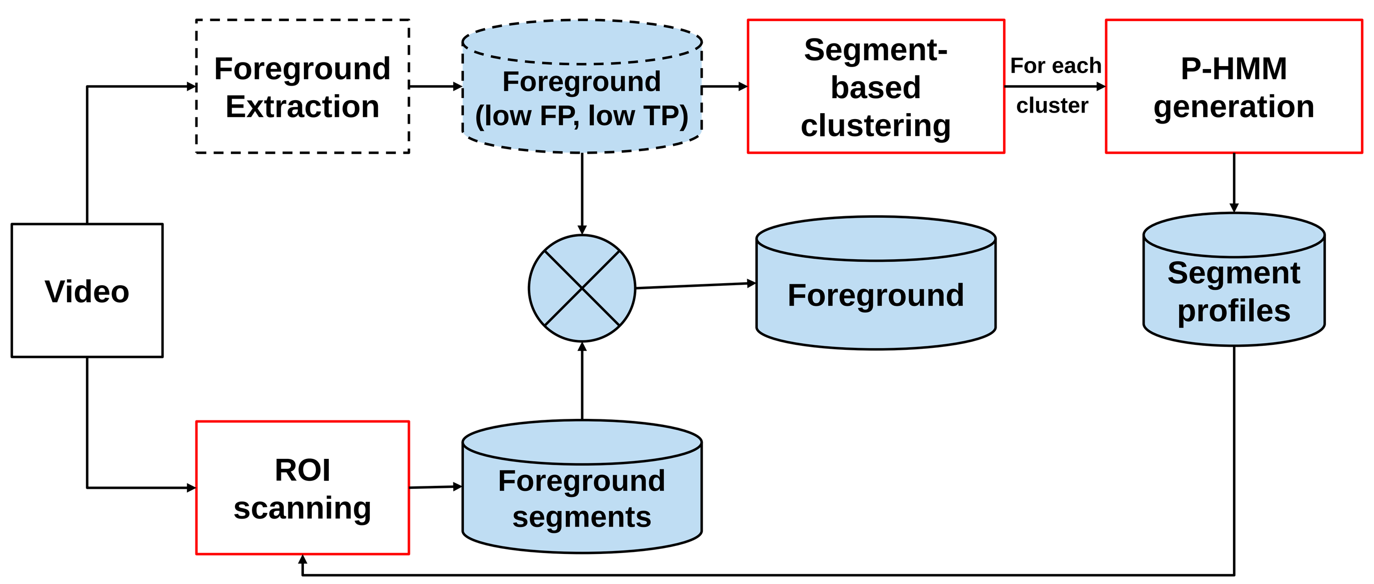
Duplication: Scene area seen by overlapping cameras

Transposition: Motion of foreground object



Foreground identification process. A foreground object, Fa, is visible in all scanlines that are analysed, occluding various background regions.

'Vide-omics'-based foreground modelling

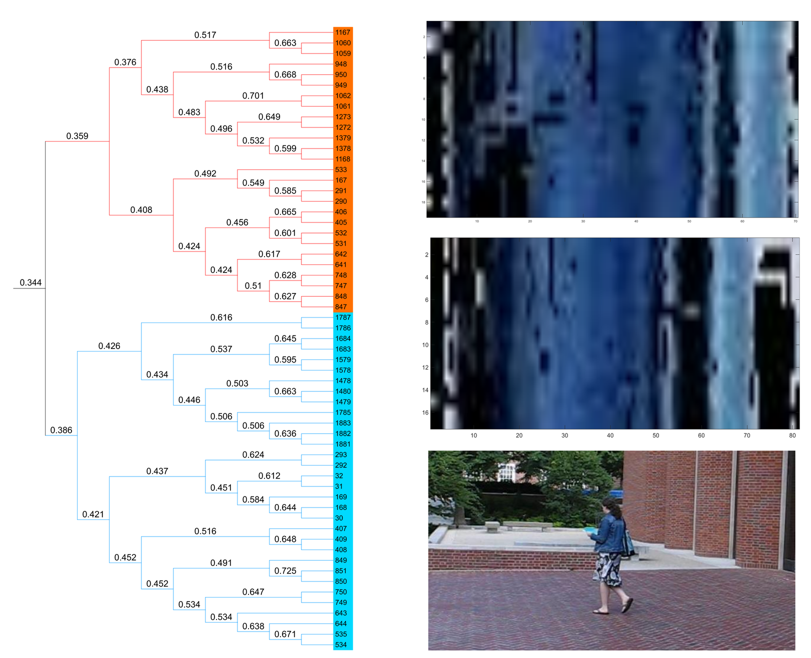


Segment-based clustering: Given a set of foreground object segments, a phylogenetic tree is built encapsulating the similarity among segments. Families of segments in the tree can be discovered by finding groups characterised by significant intra-group similarity.

P-HMM generation: For each family of segments, a profile is built encapsulating multiple correspondences across all members and it is used define the architecture, as well as, the parameters of an HMM, i.e. transition and emission probabilities.

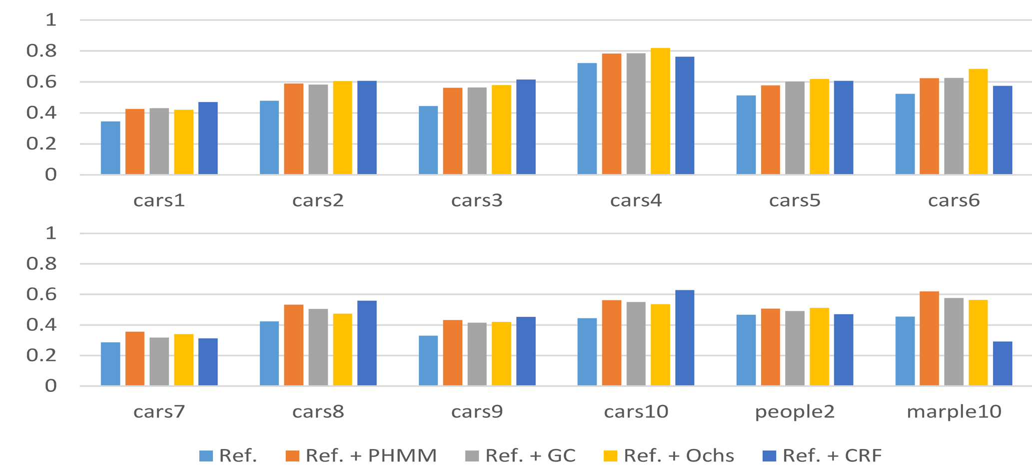
ROI scanning: Relevant segment profiles are scanned over regions of interest.

Foreground: The final foreground segmentation consists of the overlay between the initial and the detected foreground segments obtained with P-HMMs.



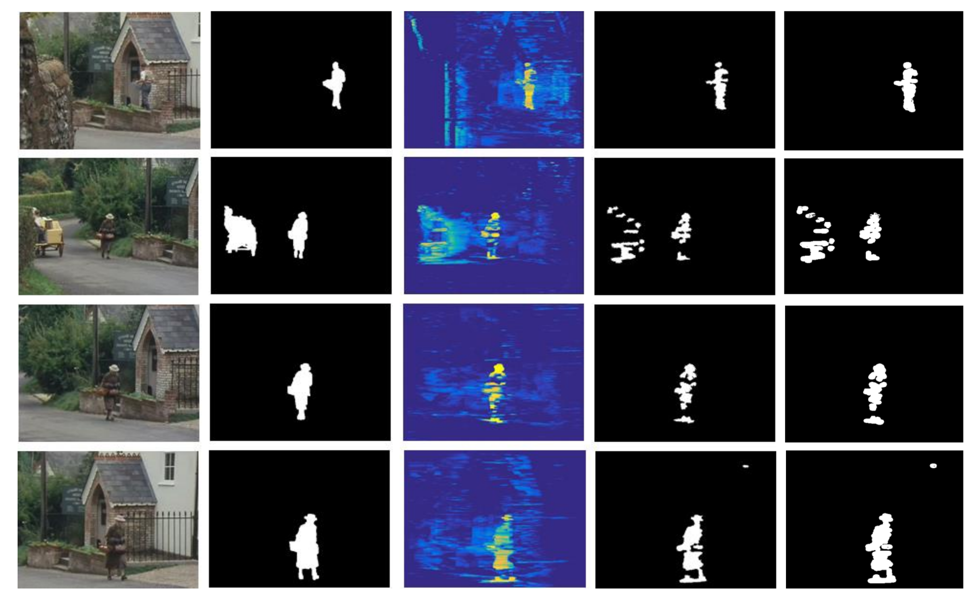
An illustration of a subtree calculated with the proposed method. In this example, the subtree consists of segments from the jeans jacket of the moving object shown above.

Results



F1 score comparison with 4 state-of-the-art methods for 12 videos.

- In terms of F1 score, the proposed method outperform significantly (~+20%) the approach used as reference.
- Both the proposed method and Ochs exhibit better performance (54.8%) than enhancements produced by either GC (53.7%) or CRF (52.9%).



Conclusion & future work

- A bioinformatics-inspired pipeline for the generation of novel object descriptors and the detection of associated objects within a video.
- Added value demonstrated by evaluation performed on a standard video dataset comprising a variety of scenes and camera motions.
- Usage of those object descriptors for direct foreground detection in unseen frames to be investigated in future work.

References

[Ka18] Kazantzidis, Ioannis, Francisco Flórez-Revuelta, Mickael Dequidt, Natasha Hill and Jean-Christophe Nebel. "Vide-omics: A genomics-inspired paradigm for video analysis." Computer Vision and Image Understanding 166 (2018): 28-40.

[Oc11] Ochs, Peter and Thomas Brox. "Object segmentation in video: A hierarchical variational approach for turning point trajectories into dense regions." 2011 International Conference on Computer Vision (2011): 1583-1590.

[Kr11] Krähenbühl, Philipp and Vladlen Koltun. "Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials." NIPS (2011).

[Ve05] Vezhnevets, Vladimir and Vadim Konouchine. "" GrowCut "-Interactive Multi-Label ND Image Segmentation By Cellular Automata." (2005).

[Gl03] Glänzel, W. & Schubert, A. Scientometrics (2003) 56: 357.