# *5C1 Video Processing : Assignment 1*

*Prof. Anil Kokaram*

*1 Jan 2023*

---

This assignment is worth 35% of your final mark. The deadline for submission of your NUKE script and short (3-5 page) report is as shown in Blackboard. PLEASE USE THE FILENAME CONVENTION `LastName_FirstName.XXX` when submitting your NUKE graph. You will be shown what data we will be using and you do not need to upload that data.

See The Foundry tutorial pages **here** for some background on NUKE graphs. See **here** for a Blink script reference.

See **here** for example kernels.

## *The task*

In lectures you worked through several levels of sophistication in designing an algorithm for image keying with colour. For this assignment you must design and implement an algorithm for binary colour keying using the data provided in lectures. Your algorithm should possess the following characteristics.

- It should implement a Bayesian interpretation of the problem *for video*.
- Use a Gaussian distribution for the background colour distribution. The parameters for that Gaussian distribution should be estimated from a block of pixels on a frame in the background.
- Use a Spatiotemporal MRF as a prior for the segmentation label field. Use a neighbourhood of $3 \times 3$ pixels spatially and $2$ frames including the current frame.
- Use motion compensation to direct the MRF along motion trajectories.
- Use ICM for solving for the segmentation mask. Use synchronous updating i.e. take a binary image as input and process that all at once to yield another binary image as output. Do not do in-place or asynchronous updating.

*The implementation*

You must implement your algorithm using NUKE scripting and node graphs. It must work for image sequences and use spatio-temporal information and not just one frame at a time. Remember that you might need to repeat nodes to achieve certain effects e.g. iterating over the same image. Use a combination of NUKE nodes and blink scripts.

*The deliverables*

The implementation must run on 5 frames input material at HD resolution. The following should be submitted for marking.

- The NUKE node graph implementing your solution.
- You will have to demonstrate your implementation running correctly to the demonstrator.
- A 3 page (min) - 5 page (max) report formatted using the IEEE style. The latex template is provided on Bb. You can see **here** for more information about the IEEE sytle itself.
- Your report should include the following information
  1. A mathematically supported description of your algorithm including references to fundamental papers/work.
  2. A description of how your algorithm maps onto your implementation
  3. An assessment of the accuracy/fidelity of your extracted mask and composite image output.
  4. An assessment of the performance impact of the following
     (a) Using 3D MRF Vs 2D MRF
     (b) Using 3D MRF versus single frame likelihood keying
     (c) Using motion versus not using motion
  5. A section on the shortcomings of your algorithm and suggestions for improvement.

*Marking Rubric*

- Language and Scientific Writing, Coherency of discussion, formatting correct : 3 marks
- Mathematical presentation of ideas correct and coherent : 4 marks
- Experiments sensible. Described cogently and succinctly : 3 marks
- Results sensible and illustrated properly using plots and images properly formatted and viewable, and referred to in the written material : 10 marks
- Nuke script works, is correct and student responds cogently to questions : 15 marks. Nuke scripts which do not work are given 0 marks.

**In the case of a submission with a NUKE script that does not work the max mark from the report will be 7 i.e. for Language and Mathematical presentation only.**

*For extra credit*

For extra credit you can include the following

1. Explore a different optimisation strategy for the labels e.g. Belief Propagation, or ad-hoc intermediate smoothing steps.
2. Explore using a larger spatio-temporal neighbourhood in the MRF.
3. Explore a non-parametric estimate for the p.d.f. of the foreground and background colour distributions.

*Tips*

The best way to do this project is one piece at a time. First get the image likelihood keyer going, then get the 2D MRF going. Then get the motion estimator in NUKE and check that it makes sense. Then connect the motion estimator and MRF module into a 3D MRF module. In the *Auxiliary Material* content area you are given test images and .nk scripts that will help you make sure that your likelihood and MRF solvers get off to a good start. Here is a suggested structure for your paper/report. You may wish to modify this as appropriate.

1. Title: Semi-automated binary segmentation with 3D MRFs
2. Abstract: Concise statement of your work.
3. Introduction: What is compositing? Why is automatic segmentation important? How is video segmentation different from image segmentation? What are the key references to Bayesian Matting and how is this different? Define your setup i.e. person against a monochromatic screen. What are your key findings? *The results show that 3D MRFs are X% bette than 2D MRFs but at the cost of Y% computational load.*
4. Video segmentation: Set up the problem by saying *Given an input image, the problem is to estimate $b(\mathbf{x})$ a binary segmentation mask which is $= 0$ in the background and $= 1$ in the foreground.* Then proceed to talk about manipulating $p(b|I, D)$ using Bayes' rule. Then propose distributions for each term .. likelihood and prior. Remember to include the penalty $\alpha$ for the likelihood when $b = 1$. Typically you'd want to have the distributions for likelihood and prior in separate subsections with a diagram for the prior.
5. Optimisation : Lay out the steps in your algorithm. Using an enumerated list. Define your terms. Talk about 2D versus 3D implementation. How do you estimate motion?
6. Implementation: Describe the nuances in your implementation e.g that you are using NUKE and blink script. What does your MRF blink script snippet look like? Maybe that could go in an appendix. Talk about 2D versus 3D implementation. How do you estimate motion?
7. Data: Show some frames from your sequence, and show the motion vectors superimposed. Use zooms. It will be good enough to show 3 zoomed in frames in a row for instance, with some motion superimposed. Show how you are going

to measure performance. One good way is to manually segment the 5 frames in the sequence and then measure how close your estimated mask is to that. Show your manually segmented masks.

8. Results: Talk about the experiments that you did and the results. Show graphs and charts to make it easier to understand your findings. For instance, plot the final error in segmentation per frame for 2D and 3D MRFS. You can use a final composite to talk about the effect of errors in the mask i.e. composite versus the background you used in class. You can also show the number of pixels marked as $1$ (say) with increasing iteration number and show how that converges rapidly .. or not. You can do the same thing with the effect of spatial smoothness.

9. Conclusions: Summarise your findings. Explain what you would do next to improve this algorithm by trying to identify at least one shortcoming and addressing it.