This section discusses the implementation of the inference models required for moving object detection. In particular, the necessary modifications to support HE video data. Adaptations were required to incorporate the HE Boolean circuits and overcome operation depth limitations. The more straightforward adjustments are summarised in §\ref{sec:stuff}, and the investigations into GMMs are given a more in-depth presentation in §\ref{sec:OMM} and §\ref{sec:EMAlg}.

# Something about converting regular algorithms to HE

# Online Mixture Model

In 1999 Stauffer and Grimson proposed \textit{adaptive background mixture models} for real-time moving object detection [STAUFFER]. To overcome a single Gaussian distribution’s inability to cope with the changing lighting conditions in practice, they proposed a mixture of adaptive Gaussians. For each frame of the video, the parameters of the Gaussians are updated, and they are heuristically evaluated to hypothesise which are most likely part of the \textit{background process}. The advantage of this technique over other GMMs is that the model runs \textit{online}. Consequently, no training phase is required. Instead, the model can be fitted, and results returned in a single phase. While this is useful for real-time inference acting on a constant stream of data, it has the disadvantage of producing less accurate results earlier in the execution sequence.

## Fitting

For a particular pixel, the values that occur over time are known as the \textit{pixel process}. This is a time series of pixel values such that, at any time $t$, the process of pixel $(x,y)$ is defined by Equation \ref{eq:pixelProcess}.

EQUATION

where $I$ is the image sequence.

Many guiding factors influence the definition of the model and updating procedure. For example, lighting changes must be tracked, static objects added to the scene must be incorporated into the background, and no camera sensor is perfect, so random noise in pixel values must be ignored. From these factors, it can be deduced that more recent observations will be more useful when determining Gaussian parameter estimates.

The recent history of a pixel can be modelled as a mixture of $K$ Gaussian distributions. $K$ is usually a value between 3 and 5, depending on available memory and computational power availability. Given a pixel process, the probability of observing the pixel value at time $t$ is given by

EQUATION

where $\omega\_{i,t}$ represents an estimate of the proportion of the data accounted for by the $i^{\text{th}}$ Gaussian at time $t$, $\mu\_{i,t}$ and $\Sigma\_{i,t}$ are the mean and covariance matrix of the $i^{\text{th}}$ Gaussian at time $t$ respectively. $\eta$ is the Gaussian probability density function given by Equation \ref{eq:pdf}.

Rather than using the \textit{expectation maximisation algorithm} (see §\ref{sec:EMAlg} for more information) to maximise the likelihood of the observed data, Stauffer and Grimson suggested using a \textit{K-means} approximation to engender the online aspect of the system. Each pixel in a new frame is compared against the existing $K$ Gaussian distributions until a \textit{match} is found. A match occurs when a pixel value is within a predefined number of standard deviations of a distribution. The number of standard deviations will vary across distributions as each distribution will account for different factors such as lighter or shadier regions.

In the event that none of the distributions match a pixel’s value, the least likely Gaussian is replaced by a new distribution defined with the pixel value as its mean, an initially high variance, and low prior weight. Then, the prior weights are adjusted at time $t$ using Equation \ref{eq:priors}.

EQUATION

where $\alpha$ is a learning rate, and $M$ is an indicator function of $1$ if Gaussian $k$ at time $t$ matched, and $0$ otherwise. After this approximation is complete, the weights are normalised.

For unmatched distributions, the $\mu$ and $\sigma$ parameters are unchanged. However, the parameters of the matching distributions are updated according to Equation \ref{eq:muAndSigma}, where $rho$ is defined by Equation \ref{eq:rho}.

EQUATIONS

## Predicting

Once the parameters have been updated, the Gaussian most likely produced by the background process must be determined to segment the foreground and background. This decision is based on the assumption that there will be a relatively little variance in the Gaussian distributions when a static, persistent object is in the frame. In contrast, when a new object occludes the background, it will generate significant variance and not match an existing distribution. Consequently, a method of defining the proportion of the GMM representing the background process is required.

To achieve this, first, the Gaussians are ordered based on the value of $\frac{\omega}{\Sigma}$. The definitions of $\omega$ and $\Sigma$ mean that this value will increase as both the distribution gains more evidence and the variance decreases. This value will only differ from the last iteration for matching distributions, so the sorting process can be made more efficient. The ordered list can then be iterated over, and the first $B$ distributions are taken as the \textit{background model}, where

EQUATION

The threshold, $T$, is a measure of how much data should be accounted for. In other words, the best-fitted distributions are taken until a certain portion of recent data has been considered.

Once the background model has been decided, it can be used to label the pixel as either \textit{foreground} or \textit{background}, allowing the moving objects to be extracted as the foreground.

# Expectation-Maximisation Algorithm