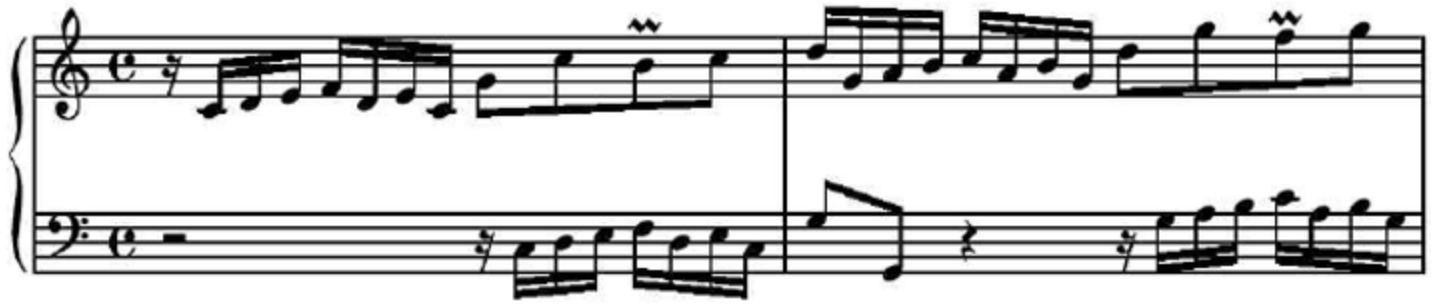
Using Mathematical Morphology for Geometric Music Retrieval

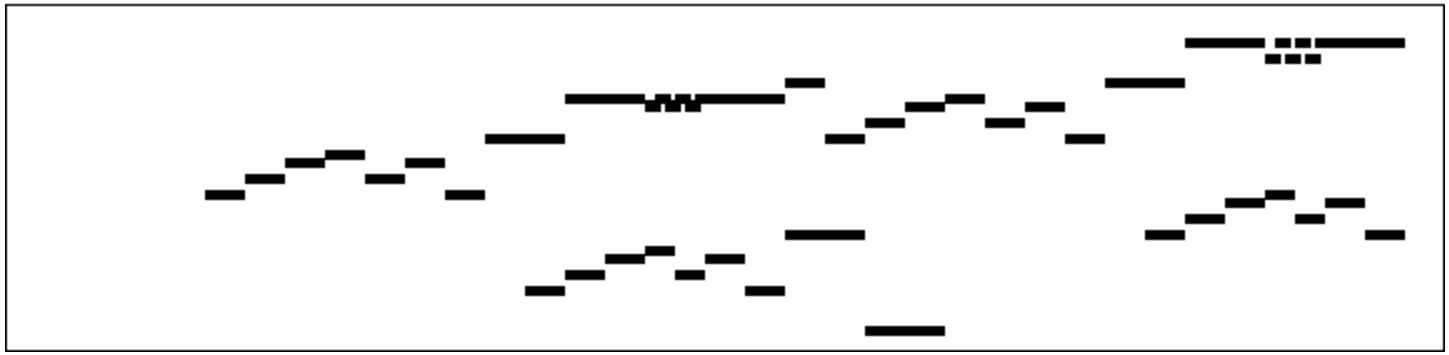
Mikko Karvonen, Kjell Lemström

University of Helsinki,
Department of Computer Science

Representation

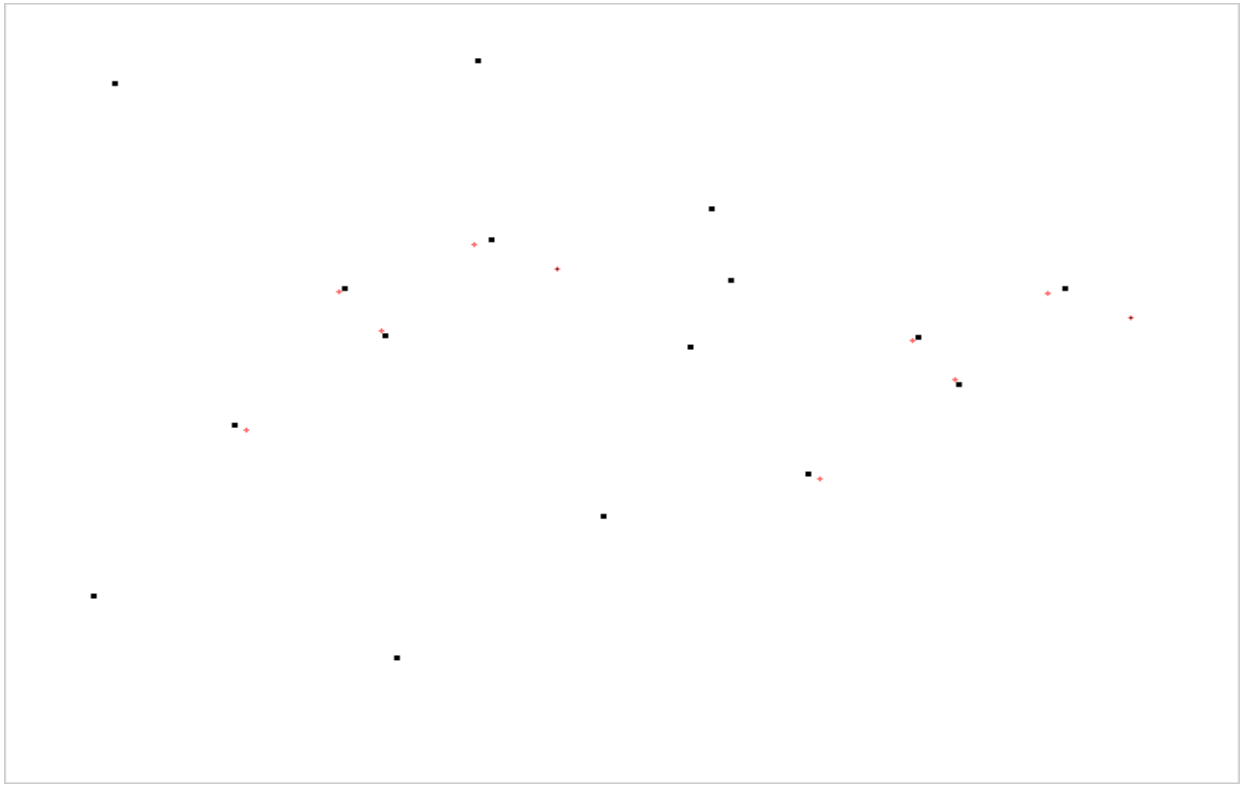
- Representation of symbolically encoded music as line-segments:





Representation

- 1-pixel representation, where the pixels of the query (marked with red dots) are slightly jittered in the database image:

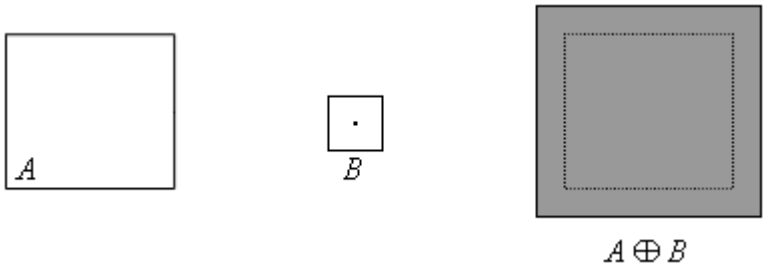


Mathematical Morphology

- A theory and technique for the analysis and processing of geometrical structures, based on set theory.
- MM was originally developed for binary images, and was later extended to greyscale and multi-band images.
- In binary morphology, an image is a subset of the integer space Z^2.
- Morphological operations are nonlinear neighbourhood operations on two sets, the typically smaller one being called *structuring element* (SE) – could be considered as a filter kernel.
- The elementary binary MM operations are *dilation* and *erosion*, based on the Minkowski addition and substraction.

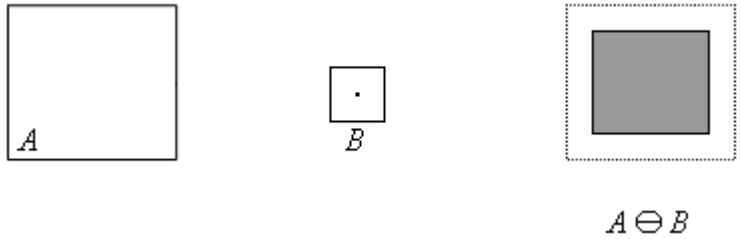
Mathematical Morphology

- Dilation performs a maximum on the SE, which has a growing effect on the target set.
- Can be used, for example, to fill gaps in an image.



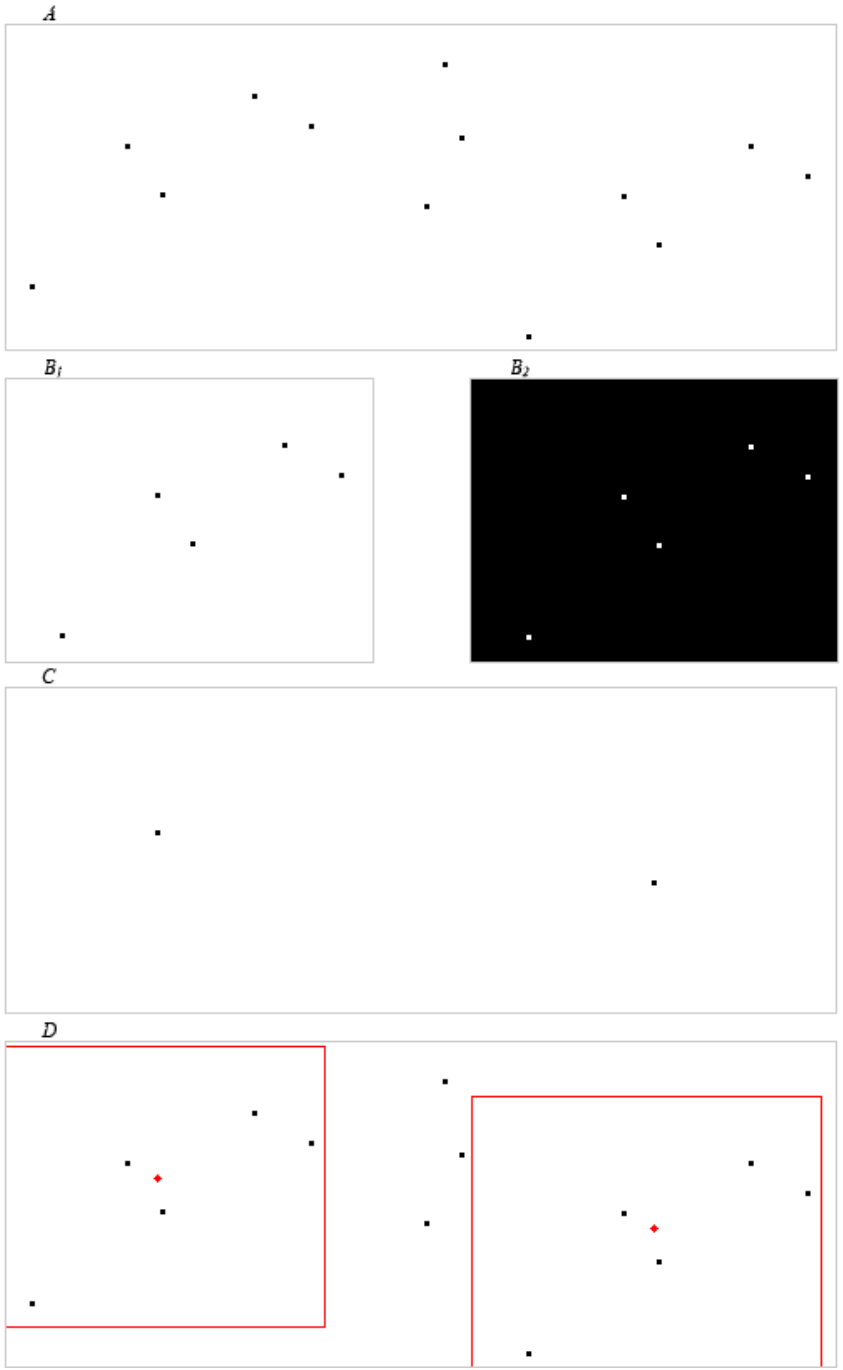
Mathematical Morphology

- Erosion performs a minimum on the SE and causes the target set to shrink.
- Can be used, for example, for removing salt-and-pepper type noise.



The Hit-or-Miss Transformation

- HMT is the intersection between an image eroded by SE_1 and the image's complement eroded by SE_2, where SE_2 is a local background of SE_1.
- The output is an image where only the occurrences are marked with an "on" pixel.

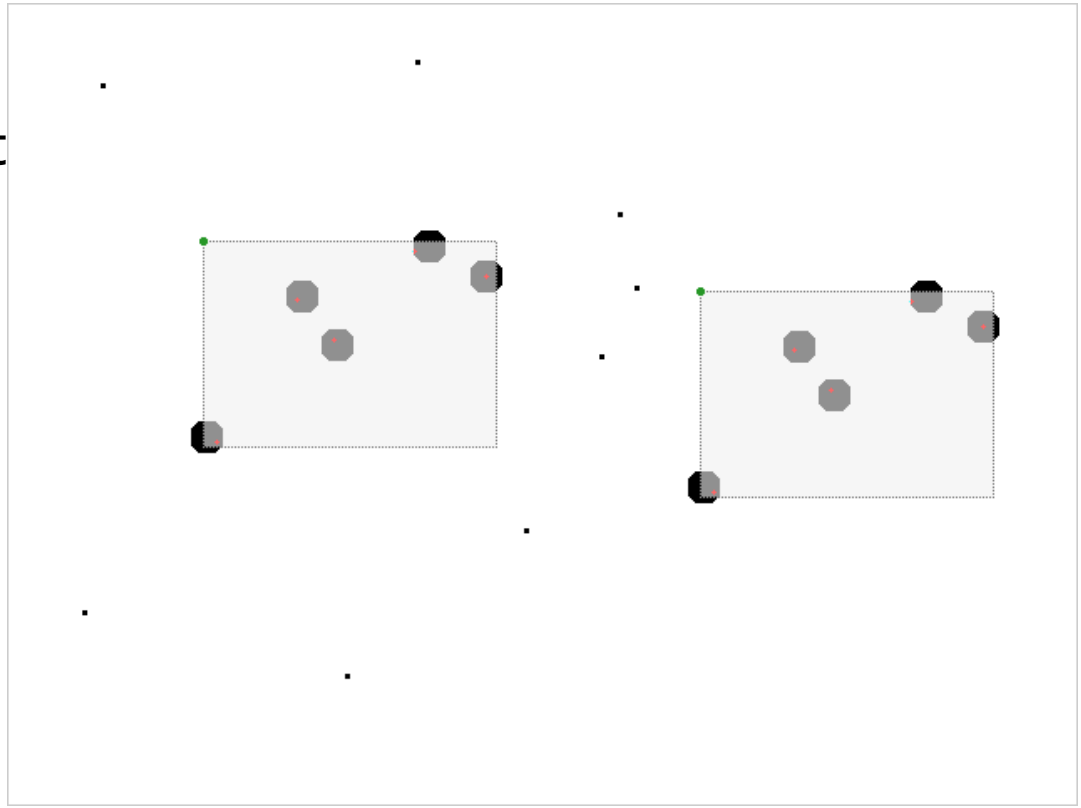


Correlation & preprocessing

- In the note-on representation, approximative matching is performed by replacing the 1-pixels in the query image by filled discs (dilation by a disc-shaped SE). This preprocessing can be done on either one of the images, but as the query is typically smaller it is faster to operate on that.
- Now, in a jittered query the pixels in the database image fall within the discs in the query.
- Intuitively, the disc allows for approximation both time- and pitchwise, and its diameter gives the magnitude of the approximation.

Correlation & preprocessing

- The red pixels mark the pixels in the database image that fall within the discs in the query.
- The top-left corner green pixels are the output of the linear correlation operation.



Threshold Convolution

- A ”modified” HMT technique.
- We define the database and the query as sets X and W, and 2-D functions $x(n)$ and $w(n)$, such that $x(n)=1$ if $n \in X$ and $x(n)=0$ if $n \in X^c$ (accordingly for $w(n)$).
- We apply a threshold value θ to the convolution of x and w, and use Heaviside’s unit step function to decide which convoluted pixels are to be used.
- If $\theta = |W|$ (the number of ”on” pixels in the query image), the operation is equivalent to the erosion of X by W.
- By lowering θ, we can impose looser conditions.

Rank Order Filtering

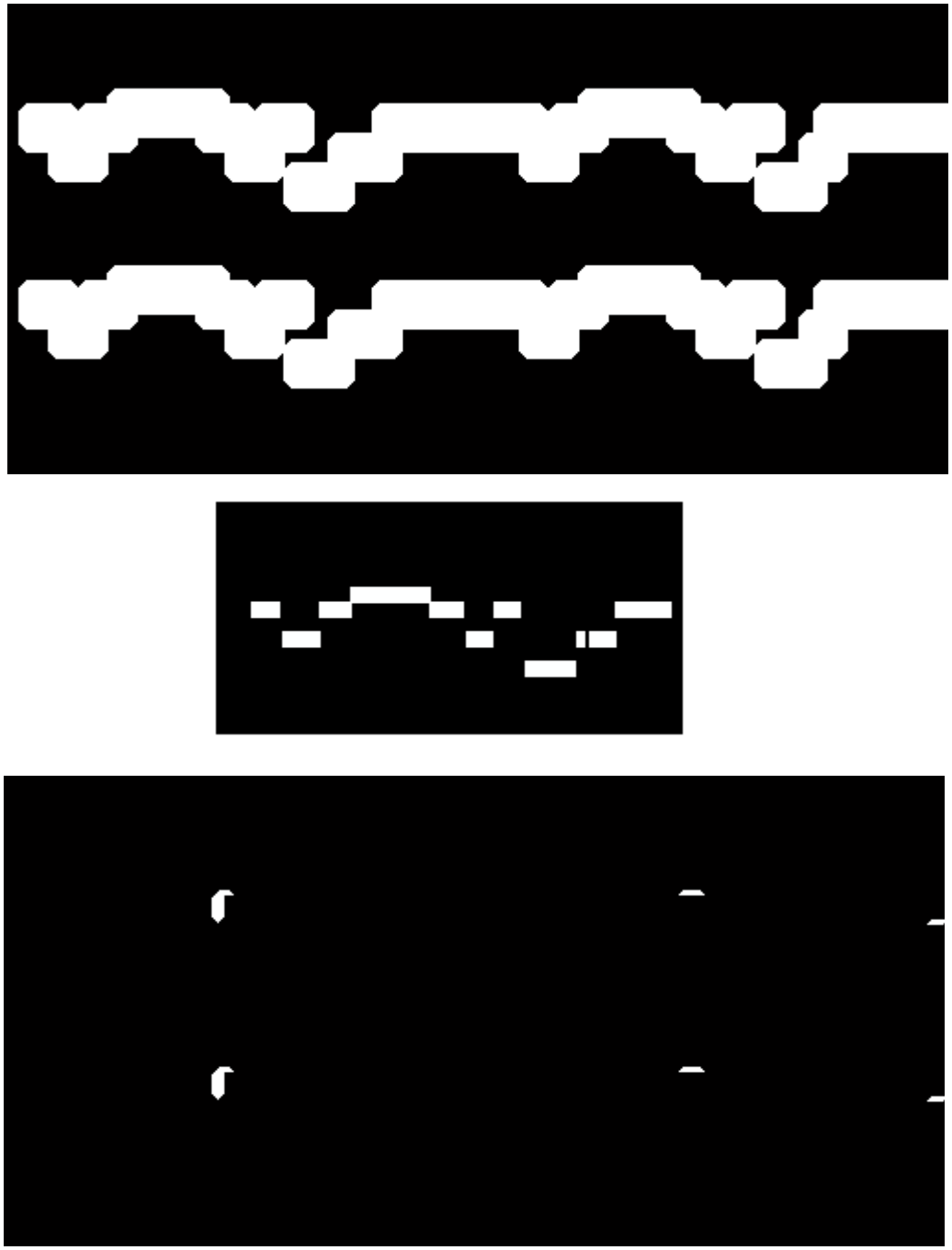
- Theoretically equivalent to TC (another ”modified HMT”.
- ROF operations are more robust against shape distortion than erosion, but they are computationally more complex.
- In the r-th rank order transformation (r=1,2,...,|W|) of the binary set X by a binary W, we shift W step-by-step. When located at point $z \in Z^2$, we always select the r-th largest of the points belonging to the intersection $X \cap (W+z)$. If r=|W|, the rank operation becomes an erosion and the filter operates as HMT. Lowering r imposes looser conditions.
- TC and ROF yield similar results when $\theta=r$.

Blur HMT

- A simple extension to the classic HMT, providing tolerance to salt-and-pepper noise and image alignment errors.
- The ”blur” parameter specifies the maximum distance allowed for a match in between the query pixel and the corresponding pixel in the database image.
- Implementable using Boolean image operators, which makes it faster than integer-based techniques, such as ROF.
- The only change when compared to the original HMT is the inclusion of a straight-forward preprocessing: the image is first dilated by a disc-shaped SE_1 and then the image’s complement by another disc, SE_2.

Blur HMT

- Top: the database image dilated by a disc-shaped SE (of 10-pixel radius).
- Middle: the query image.
- Bottom: the output of the Blur HMT marks the near-matches.



Conclusions

- Each of our 4 algorithm prototypes based on MM are capable of finding occurrences of query patterns that had minor distortions.
- In note-on representation, with a small database image, ROF was the fastest; with larger databases TC outperformed the others.
- We expect BHMT to be the choice when implemented by using Boolean image operations.
- In line segment representation, linear correlation with morphological preprocessing was the fastest approach, but it also gave false positives most often.