Streaming Thin Client Compression

- Bernd Oliver Christiansen
- Klaus Erik Schauser
- Malte Muenke

- University of California

Thin Client

- Model in which applications are executed on a server and only input events and screen updates travel the network.
- Requires fast and efficient (de)compression of synthetic images

Synthetic Images

- Fewer colors than pixels
- Sharply delineated uniform-color regions with overlaid symbols
- Do not contain regions with smoothly varying pixel intensities
- Contain redundant features/blocks
- Do not compress well with lossy methods. Text can become unreadable

Previous Work

- GIF (Graphics Interchange Format)
- PNG (Portable Network Graphic)
- FABD (Flexible Automatic Block Decomposition)
- PWC (Piecewise-constant)

PWC

- Four questions:
  - Q1: Is the current pixel the same value as the one to the left or above?
  - Q2: Is the current pixel the same value as its top-left or top-right neighbor?
  - Q3: Does the current pixel equal a guess? (uses known surrounding colors for probability estimation)
  - Q4: What is the pixel's value?

PWC cont’d…

- Decoder “asks” the questions in order.
- Different depths of coding can be used depending on the complexity of color.

TCC

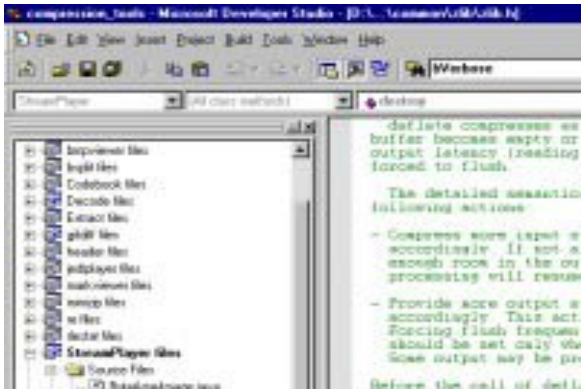
- Basic idea
  - Scan the image for marks
  - Build a codebook (dictionary) of marks
  - Replace marks by dictionary references
  - Code the residue and codebook with PWC

Segmenting the Image into Marks

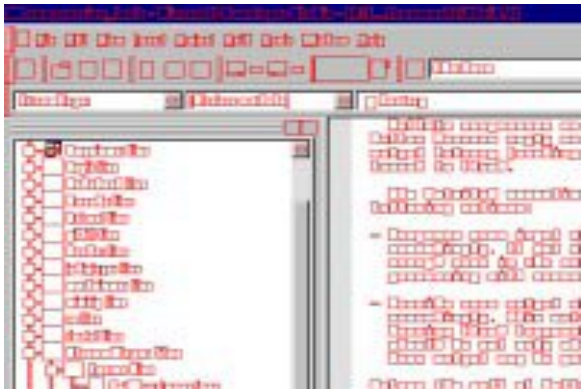
- Mark: a set of connected pixels that is surrounded by a single-color boundary where no pixel in the set is the same color as the boundary
- Any color transition indicates a “seed”
  - Walk “with the right hand on the wall” along pixels of the same color as the seed

Marks cont'd…

- Exact shape is not extracted. A bounding box is acquired.
- After saving in codebook, the bounding box is filled in with the seeds color.
- Marks are empirically observed to be small.
  - Bounding box restricted to 48x48 pixels maximum



Original



Marks

Codebook

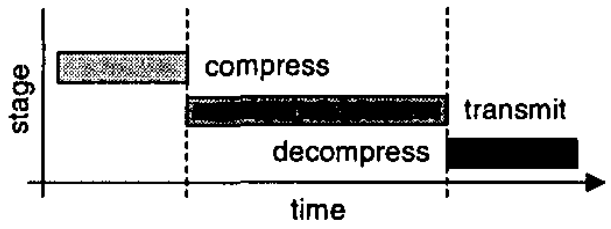
Background

Codebook

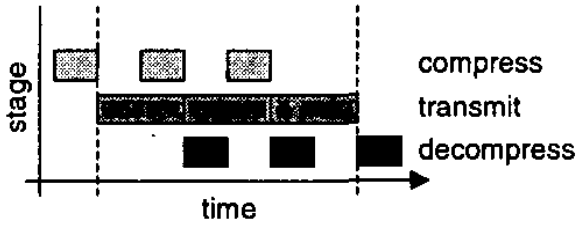
- As each bounding box is extracted the library is searched for an *exact* match.
- In Codebook marks are coded in PWC as well as width and height.
  - Bilevel marks are coded in depth 1 PWC
  - Non Bilevel in depth 4 PWC

Streaming Thin Client Compression

- End to end latency is most important consideration for performance of Thin Clients
- The Streaming version of TCC addresses latency by modifying TCC to allow pipelining.



(a) Sequential Transmission



(b) Pipelining

Streaming cont'd…

- Original TCC requires two passes.
  - One to extract marks and create codebook
  - A second to code the residue
- STCC makes only one pass and fully compresses each row which can then be sent in the pipeline.

Streaming cont’d…

- Issues with only one pass:
  - Must trace in parallel the contours of all marks in the current row.
  - Streaming requires each row to be encoded immediately, codebook must support incremental pointers to partial marks.

Streaming Boundary Tracing

- L and R brackets
  - Do not know if a mark is connected in a following row.
  - Does not trace exact contour of the mark.

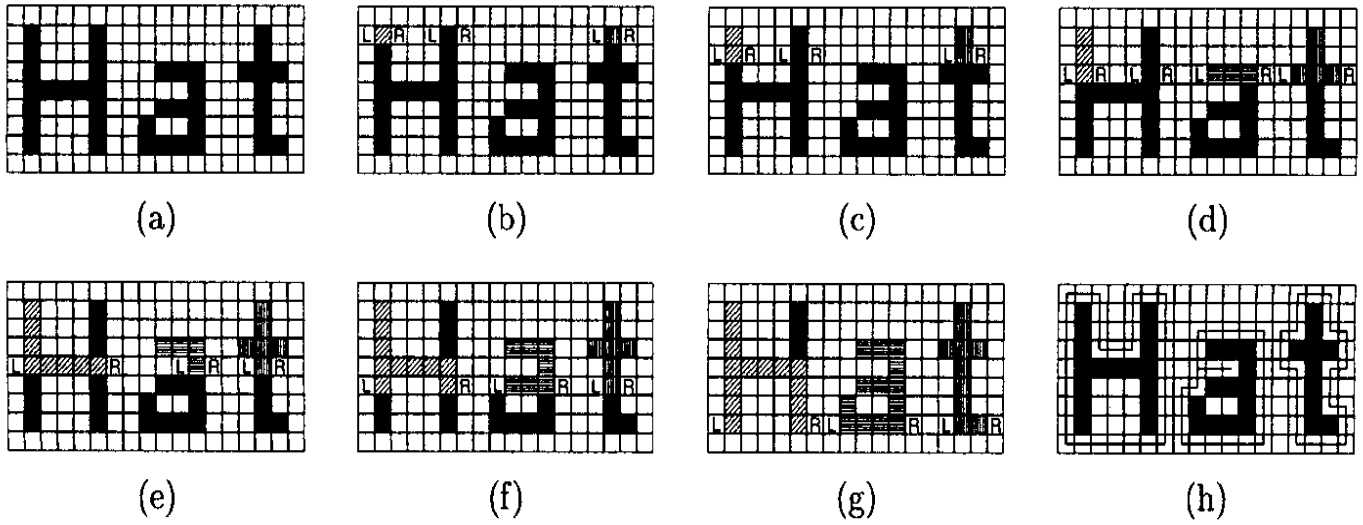


Figure 2: *Streaming boundary tracing with L-R brackets.*

Tree Structured Codebook

- Marks are stored as a path whose nodes are labeled by the mark's pixels.

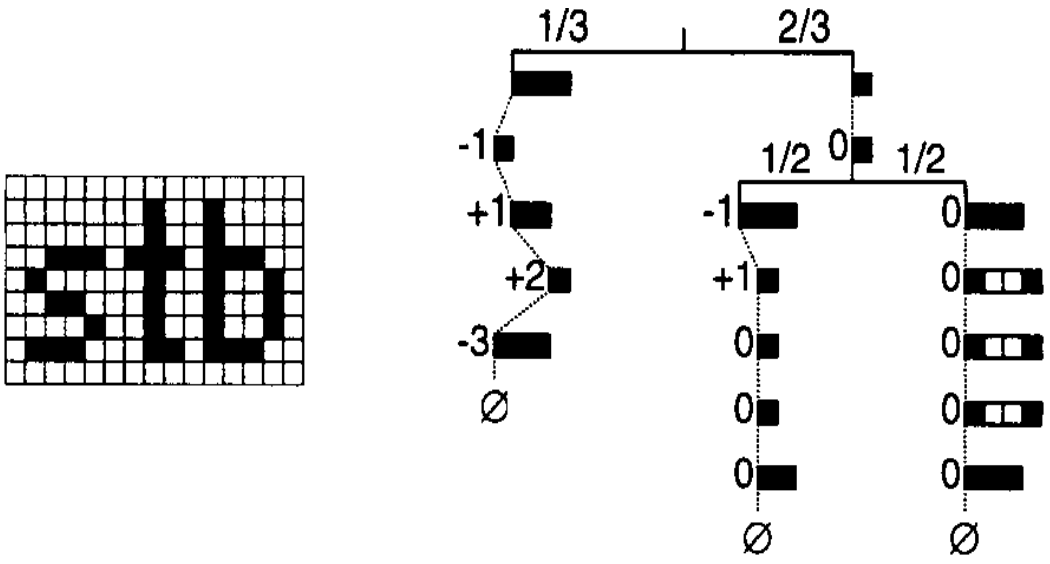


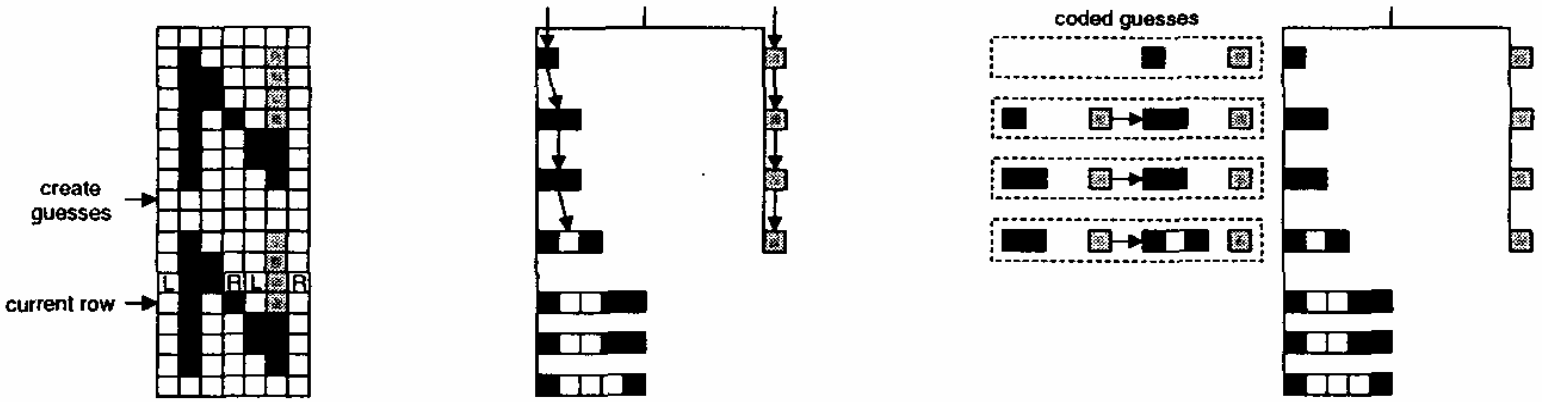
Figure 3: *A bitmap and the corresponding tree–structured codebook.*

Shortcomings of Tree Codebook

- Most nodes in the codebook only store pixels and do not have siblings
- When representing the codebook itself assume a high probability for single child nodes and have an escape sequence for branching nodes.

Shortcomings cont'd…

- Guessing connected components
  - When a component that started out as two marks is found to be connected employ a mechanism to guess the connection in the future.



(a) Image (b) Tree-based coding (c) Guess-based coding

Figure 4: *Coding (partial) marks of a connected component as a unit.*

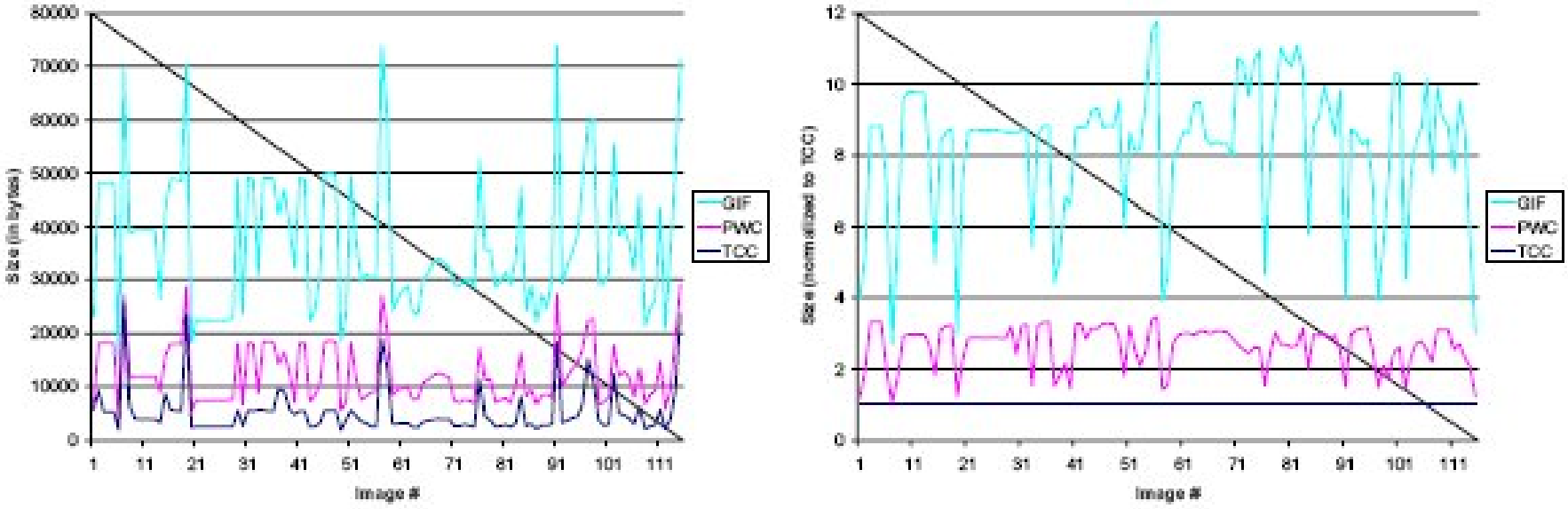
Shortcomings cont'd…

- STCC identifies more marks than TCC due to a less strict definition of marks.
  - Increased size of codebook increases pointer size.
  - Maintain a separate codebook tree for each background color. Decoder looks at appropriate tree given the background color.

Conclusion

- PWC is very good at coding uniform blocks of color
- TCC extracts marks into a codebook allowing PWC to perform as best as possible
- STCC pipelines TCC to reduce latency

Results TCC



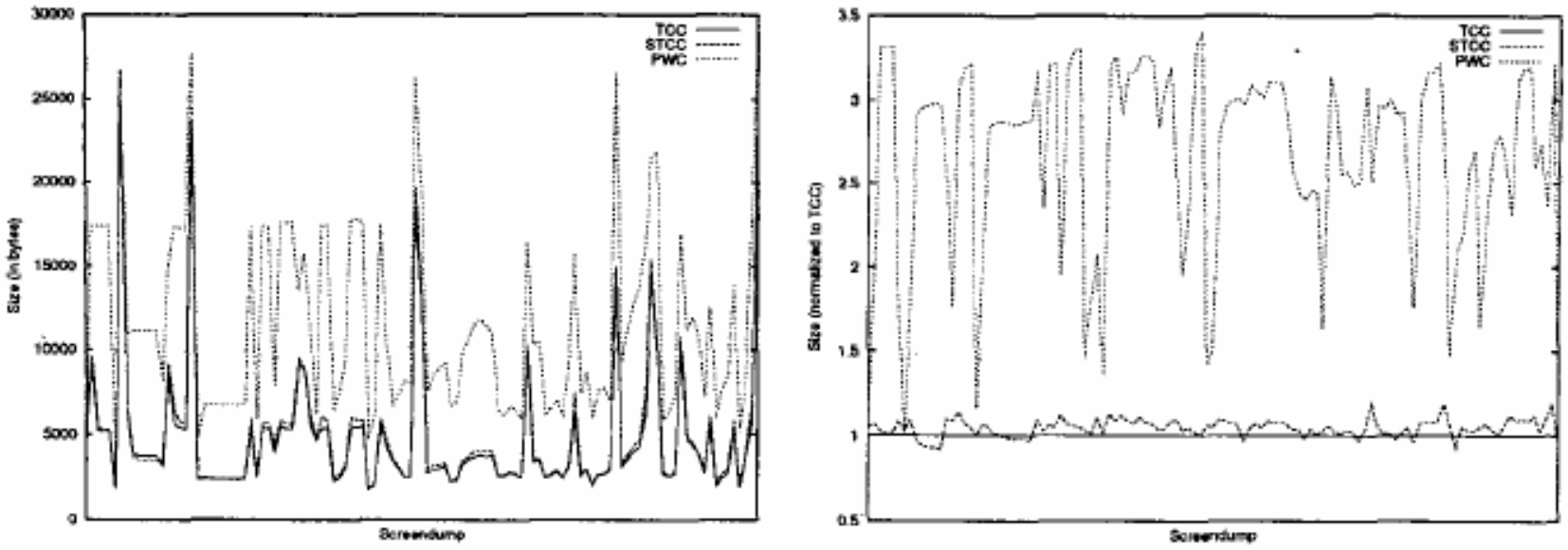
(a) in bytes

(b) normalized to TCC

Figure 5: *Compressed size of a series of screendumps. The screendumps were taken every five minutes on a Windows NT workstation running at a resolution of 1152x864 and then quantized to an eight bit palette. They show mostly Outlook and various web pages within IE.*

Spikes are due to images in web browser pages

Results STCC

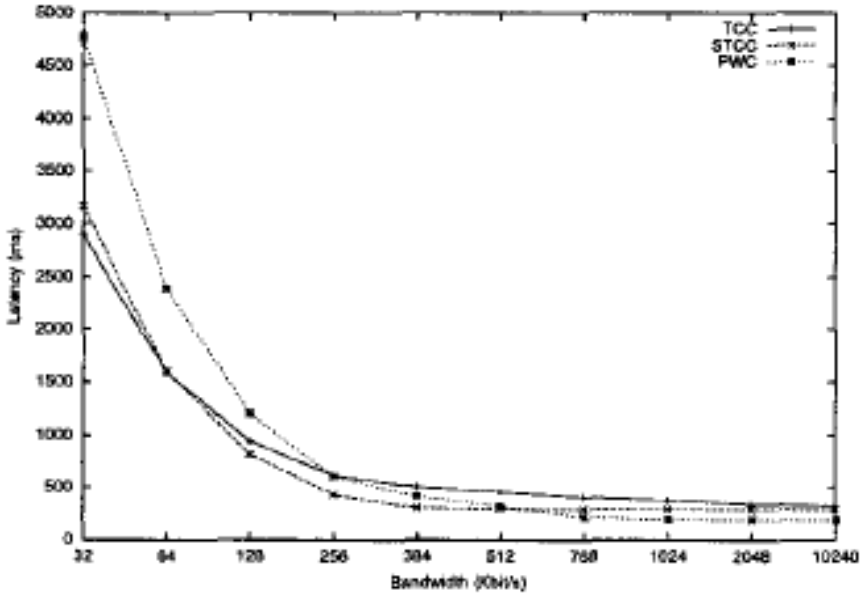


(a) in bytes

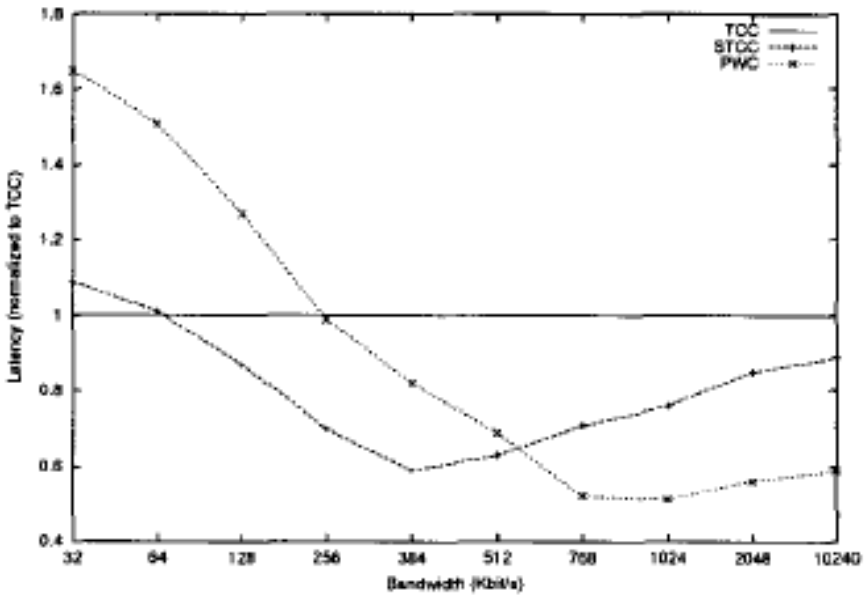
(b) normalized to TCC

Figure 6: *Size of 115 screendumps that were taken over a day.*

Results STCC cont'd…



(a) in milliseconds



(b) normalized to TCC