Efficient Clustering-Based Algorithm for Predicting File Size and Structural Similarity of Transcoded JPEG Images

Steven Pigeon
Stéphane Coulombe

Efficient Clustering-Based Algorithm for Predicting File Size and Structural Similarity of Transcoded JPEG Images

Steven Pigeon & Stéphane Coulombe
pigeon@iro.umontreal.ca
stephane.coulombe@etsmtl.ca

Dept. of Software and IT Engineering
École de Technologie Supérieure
1100 Notre-Dame Ouest, Montréal
Québec, Canada

December 2, 2011

Overview of the presentation

We will show that the proposed prediction method, EJQSP, for JPEG adaptation outperforms significantly our previous work.

The presentation is structured as follows:

Problem Definition

Adapting Images to Given Constraints

Why Adapt Images?

- Different contexts: MMS
- Different contexts: Universal Access

Problem Context

MMS — *Dramatis Personæ*

Alice

Minou

Alice's phone has a
quite capable phone:

10 MPixel Camera,
Plays media like H.264,
3G, E-Mail, etc.

On the other hand,
Bob's phone is a
quite limited phone:

Does not have a camera
Pictures upto 640x480
Messages up to 100K

Bob

Problem Context

MMS — Heterogeneous Terminals

$\leftarrow$ COMPATIBLE $\leftarrow$

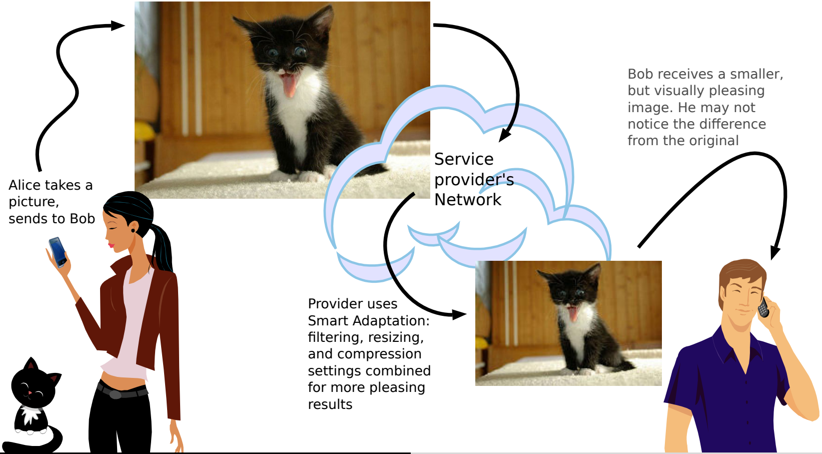
$\rightarrow$ INCOMPATIBLE! $\rightarrow$

E-mail, "Megapixel" Profile

"Image Rich" profile

Problem Context

MMS — User-Experience Based Adaptation



Problem Context

Mobile Browsing

Mobile Browsing

- **Many different devices**
  that are not all very capable
- **Different bandwidths**
  that depends on data plans
- **Inherently richer than MMS**
  Images of all types, video, scripts, etc.
- **Different goals**
  like maximizing responsiveness while minimizing battery usage

Problem Definition

In order to optimize content for MMS or for mobile browsing, we need:

- A domain-specific model of user-experience

...and constraints:

- Maximum resolution of images,
- Maximum message size,
- Image and Video coding standards,
- Transmission Bandwidth,
- etc.

Adaptation

Harder than it Looks

This *seems* trivial, but...

- Adaptation is useful under explicit optimization
  The goal is to adapt document (MMS, Web Page) in a way that maximizes user-experience under the given constraints (bandwidth, resolution, message size), and as such,
- Adapting JPEG efficiently is still a challenge
- Previously proposed solutions require partial decompression
  Compressed-domain solutions are also very restrictive, e.g., resizing by powers of 2 (and still compute-intensive)
- Adaptation must be machine-efficient
  Adaptation must be performed extremely fast to accommodate a large number of messages/pages.

Proposed Solution(s)

In previous work, we have proposed to guide explicit optimization of MMS messages [4] using predictors [1–3]

The rationale being that low cost predictors avoid performing actual transcodings until the best (probable) solution is found by optimization.

In this paper, we propose a new predictor based on clustering to guide the efficient adaptation of JPEG images.

Transcoding Operations

A transcoding operation is defined as the output quality factor, QF_{out} and scaling $0 < z \leqslant 1$, to apply to an image with original resolution $w \times h$ and quality factor QF_{in}.

The transcoding operation (QF_{out}, z) applied to an image I yields

- a quality $q(I, QF_{out}, z)$, measured using SSIM
- and a file size $f(I, QF_{out}, z)$.

...which we want to predict without actually performing the transcoding.

JQSP1

JPEG Quality and Size Predictor

Predicts the file size and quality resulting from a transcoding operation (new scaling, z, and quality factor QF_{out}), given the input image file size and original quality factor QF_{in} [1,2].

The parameters where quantized for prediction: we have $\widetilde{z}$, $\widetilde{QF}_{in}$, $\widetilde{QF}_{out}$, indexing an array which contained the predictions for the resulting file size and for the resulting quality.

The prediction for quality (or relative file size) is the expectation of quality (or relative file size) over all transcodings with the same parameters, estimated over a (large) corpus of images.

JQSP1

- The quality factors $\widetilde{QF} \in \{10, 20, \ldots, 100\}$
- The scalings $\widetilde{z} \in \{0.1, 0.2, \ldots, 1.0\}$.
- $\sim 73\,000$ original images
- For each image, 100 different transcodings: all $(\widetilde{QF}, \widetilde{z})$
- $\sim 6\,570\,000$ training exemplars (leaving out ~ 730000 test exemplars in a 90%/10% scheme).
- This corpus was used to populate a $10 \times 10 \times 10$ (we also have $\widetilde{QF}_{in}$) array for predictions $\hat{f}(I, \widetilde{QF}_{out}, \widetilde{z})$ and $\hat{q}(I, \widetilde{QF}_{out}, \widetilde{z})$

JQSP2

Predictor JQSP2 [3] does the converse of JQSP1 in that it predicts the transcoding parameters that maximize perceived quality given a file size constraint.
(rather than directly predicting the resulting file size and quality)

As JQSP1, JQSP2 uses the expectation to formulate its predictions, but does not use quantization directly: it will use all training exemplars that minimize the error to the desired file size while maximizing the quality.

Proposed Solution

In this work, we propose a new solution, EJQSP, based on clustering.

The general idea is to represent transcoding operations on images as points in a (moderately-)high dimensional space and to partition this space in order to make the predictions.

The Clustering Problem

Given $X = \{x_1, x_2, \ldots, x_n\}$, n exemplars in $\mathbb{R}^d$,

We define a partition $C = \{C_1, C_2, \ldots, C_m\}$ such that $\bigcup_{i=1}^{m} C_i = X$ and $C_i \cap C_j = \varnothing$ for $1 \leqslant i \neq j \leqslant m$, and that for each C_i, we have a centroid given by

$$\bar{x}_i = \frac{1}{|C_i|} \sum_{x_j \in C_i} x_j$$

The error associated to the partition C is therefore:

$$E(C) = \sum_{i=1}^{m} \sum_{x_j \in C_i} \|x_j - \bar{x}_i\|^2$$

for an appropriately defined metric $\|\cdot\|$

The Clustering Problem

The goal is to find the optimal partition C^*, such that

$$C^* = \arg\min_C E(C) \tag{1}$$

But this is NP-hard, except for trivial cases!

We will use K-Means to solve eq. (1)

Proposed Solution

Exemplars

For our problem, we will define each of the exemplars as a 9-dimensional vector given by:

$$x_j = (QF_j, w_j, h_j, b_j, QF_{out}, z, QF_{out} - QF_j, f, q) \qquad (2)$$

where

- QF_j is the original quality
- w_j and h_j are the normalized width and height,
- b_j is the bits per pixel of the image,
- QF_{out}, the desired output quality,
- z, the scaling,
- $QF_{out} - QF_{in}$ a "feature",
- f and q are the resulting file size and quality, by abuse of notation.

Proposed Solution

Prediction

For a new image I_j, to transcode using transcoding parameters QF_{out} and z, we form a vector such as given by eq. (2), and we find the closest centroid's index given by:

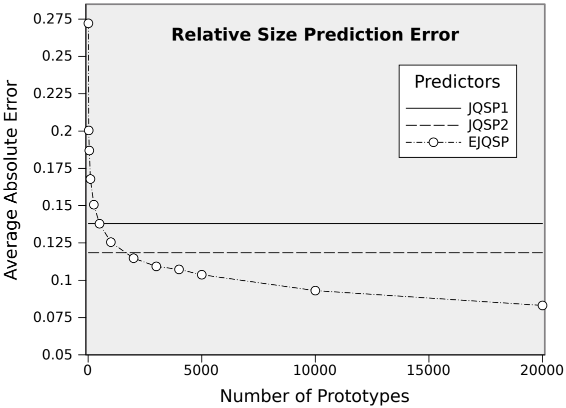
$$i = \arg\min_i \|x_j - \bar{x}_i\|^2$$

where the f and q components are left out of the metric—as they are the quantities we want to predict!

The predictions are thus: $\hat{f} = \bar{f}_i$, the f-component of $\bar{x}_i$, and $\hat{q} = \bar{q}_i$, the q-component of $\bar{x}_i$.

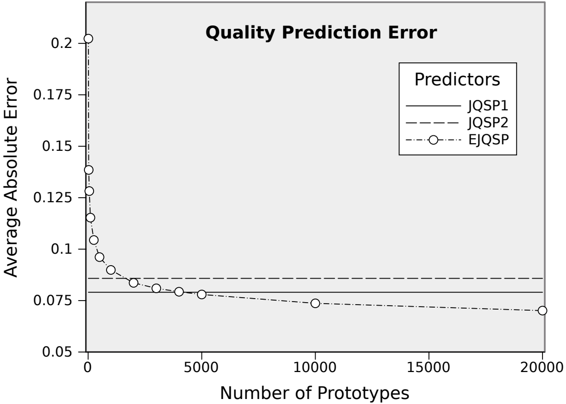
Results

Relative File Size Prediction Error



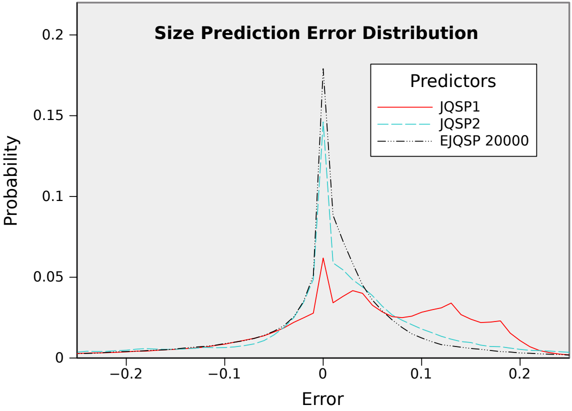
Results

Quality Prediction Error



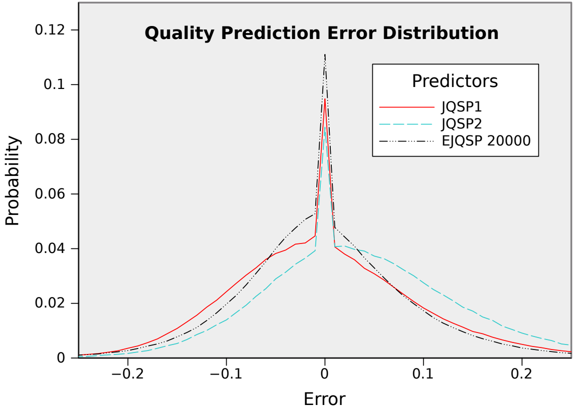
Results

Relative File Size Prediction Error Distribution



Results

Quality Prediction Error Prediction



Results

summary

EJQSP bests previously proposed method significantly:

- EJQSP File size prediction error:
  - $\approx 40\%$ smaller than JQSP1
  - $\approx 27\%$ smaller than JQSP2
- EJQSP Quality Prediction error:
  - $\approx 20\%$ smaller than JQSP1
  - $\approx 12\%$ smaller than JQSP2

Training

All (initial) training is done off-line on the exemplar from the image corpus.

- For JQSP1, training (once the exemplars are obtained), is essentially $O(n)$
- For JQSP2, training is $O(n\,|\widetilde{QF}_{out}|\,|\tilde{z}|\,|\tilde{f}\,|)$
- For EJQSP, training is $O(d\;\; m\;\; n \log n)$, for m prototypes from n exemplars in $\mathbb{R}^d$

Prediction

For prediction, we have that:

- JQSP1 predicts in $O(1)$,
- JQSP2 predicts in $O(1)$,
- ...but EJQSP predicts in $O(m\ d)$, because nearest neighbor in $\mathbb{R}^d$ is hard

Conclusion

In summary:

- EJQSP predicts much better file size and quality than our previous work
- EJQSP prediction has higher algorithmic complexity, but still reasonable
- EJQSP can be used in optimization systems such as presented in [2,3]

References

[1] Steven Pigeon, Stéphane Coulombe — *Computationally Efficient Algorithms for Predicting the File Size of JPEG Images Subject to Changes of Quality Factor and Scaling* — In Procs. 24th Queen's University Biennial Symposium on Communications, 2008.

[2] Stéphane Coulombe, Steven Pigeon — *Quality-Aware Selection of Quality Factor and Scaling Parameters in JPEG Image Transcoding* — In Procs. IEEE 2009 Computational Intelligence for Multimedia, Signal, and Video Processing (CIMSVP)

[3] Stéphane Coulombe, Steven Pigeon — *Low-Complexity Transcoding of JPEG Images with Near-Optimal Quality Using a Predictive Quality Factor and Scaling Parameters* — IEEE Trans. Image Processing, 19(3):712–721, March 2010

[4] Steven Pigeon, Stéphane Coulombe — *Optimal Quality-Aware Predictor-Based Adaptation of Multimedia Messages* — Procs. Intelligent Data Acquisition and Advanced Computing Systems (IDAACS) 2011 (september 15-17 2011) V1 p. 496-499

This work was sponsored by

NSERC CRSNG

http://www.nserc-crsng.gc.ca/

vantrix

http://www.vantrix.com/