ARTICLE IN PRESS

Available online at www.sciencedirect.com
ScienceDirect

Pattern Recognition ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Real-time detection of steam in video images

R.J. Ferrari[a,*], H. Zhang[a], C.R. Kube[a,b]

[a]*CIMS Laboratory, Department of Computing Science, University of Alberta, Edmonton, AB, Canada T6G 2P8*
[b]*Research Department, Syncrude Canada Ltd., Edmonton, AB, Canada T6N 1H4*

Received 21 December 2005; received in revised form 7 July 2006; accepted 12 July 2006

Abstract

In this paper, we present a real-time image processing technique for the detection of steam in video images. The assumption made is that the presence of steam acts as a blurring process, which changes the local texture pattern of an image while reducing the amount of details. The problem of detecting steam is treated as a supervised pattern recognition problem. A statistical hidden Markov tree (HMT) model derived from the coefficients of the dual-tree complex wavelet transform (DT-CWT) in small 48×48 local regions of the image frames is used to characterize the steam texture pattern. The parameters of the HMT model are used as an input feature vector to a support vector machine (SVM) technique, specially tailored for this purpose. By detecting and determining the total area covered by steam in a video frame, a computerized image processing system can automatically decide if the frame can be used for further analysis. The proposed method was quantitatively evaluated by using a labelled image data set with video frames sampled from a real oil sand video stream. The classification results were 90% correct when compared to human labelled image frames. The technique is useful as a pre-processing step in automated image processing systems.

Keywords: Steam detection; Video segmentation; Detection of dynamic texture; Oil sand

1. Introduction

Alberta's oil sands industry produces over one million barrels of oil per day, which accounts for approximately 33% of Canada's total oil production. Oil sand is first mined, crushed and then screened before hot water is added to extract the oil. Real-time measurements from video images of oil sand fragmentation can be used to improve the crushing and screening process. However, the intermittent presence of steam in the images, due to the cold ambient temperatures found in Northern Alberta, can affect measurements made from video images using typical image processing algorithms.

Crushed oil sand is composed of sand, bitumen, mineral-rich clays, water and comes in a variety of shapes, sizes, colors and textures [1]. Truck and shovel operations mine the oil sand, which is then crushed in double roll crushers, screened for oversize fragments and mixed with hot water producing a slurry that is then sent for further processing in an oil extraction facility. Video cameras, installed at key locations in the size reduction process, provide real-time images available for size analysis under varying outdoor mining conditions. Oil sand fragment size data can be used to monitor and improve the performance of the crushing and screening equipment. Real-time automated image processing systems, used in size analysis, depend on good quality high contrast images in order to correctly segment and measure oil sand fragment size, including oversize lumps [2].

The intermittent and unpredictable presence of steam can affect image quality and the results produced by automated image processing systems. Ambient temperatures in Northern Alberta can vary from minus to plus 40 °C, which when combined with hot water and an outdoor mining environment can produce video segments with varying amounts of steam. In this context, detection and quantification of

* Correspondent author. Tel.: +1 416 987 7528.
E-mail addresses: ricardo.jose.ferrari@gmail.com, ferrari@cs.ualberta.ca (R.J. Ferrari), zhang@cs.ualberta.ca (H. Zhang), kube@cs.ualberta.ca (C.R. Kube).

0031-3203/$30.00
doi:10.1016/j.patcog.2006.07.007

regions within the video frame that are covered by steam is an important task, and the result can then be used to determine if a video frame is suitable for further processing.

In this paper, we present a real-time image processing technique for the detection of steam in oil sand video images. The problem of detecting steam is treated as a supervised pattern recognition problem which uses the dual-tree complex wavelet transform (DT-CWT) and the statistical HMT model computed for small 48×48 local regions of the image frames in order to characterize the steam texture pattern. By detecting and providing the total area covered by steam in a video frame, a computerized image processing system can automatically decide if the frame can be used for further analysis.

The rest of this paper is organized as follows: Section 2 describes previous research on detection of steam and other closely related fields. The methodology used in this work is presented in Section 3. The experimental results and discussions are described in Section 4, followed by conclusions in Section 5.

2. Previous research

Several process monitoring problems within the oil sands industry are potentially solvable using computer vision techniques. Computer vision techniques (i.e., edge-based), previously developed for measuring ore-size distribution in the hard-rock mining industry [2], have difficulty segmenting images of oil sand due in part to the light/dark oil sand composition [3–5]. Computerized detection of steam, for instance, is a very specific and challenging problem in monitoring ore-size distribution in oil sands industry. Besides the importance in the oil sands industry, an automatic algorithm for detection of steam could also be adapted to other industrial applications such as fire alarms, cargo smoke detection, and other video smoke detection systems.

In Ref. [6], for instance, after a simple background subtraction procedure, the authors used the grayish color information, indicated by the intensity of the HIS color model, to detect smoke pixels in an image. A decision function is built from a set of threshold values derived from different smoke scenes. However, as pointed out by the authors, the problem with chromatic models is that changes in the fuel type may affect considerably the computed threshold values. In Ref. [7], presence of smoke in the field-of-view of the camera is detected by monitoring the decrease in the energy content of the original scene and in the estimated background using the discrete wavelet transform (DWT). In addition, the authors use the flicker frequency of the flames to identify possible steam boundaries, and a decrease in chrominance value of the pixels. The steam region is also considered as convex region by the authors.

In Refs. [8,9], detection of smoke is treated as a dynamic texture problem. The authors proposed to model the spatio-temporal dynamics of image regions by using Gauss–Markov models, and infer the model parameters as well as the boundary of the regions in a variational optimization framework using the level-set technique.

A few systems for detection of fire and smoke are also commercially available [10,11]. However, for obvious reasons, they do not provide enough details about the algorithms.

Despite the previously described research projects, the problem in detecting steam in oil sand video images is quite unique due to the fact that the image background is not static, making background subtraction and optical-flow techniques not helpful in detecting the steam boundaries on an image. Also, changes in lighting (from clouds, time of day, etc.) is another factor that can easily disturb these methods.

Inspired by the work of Töreyin et al. [7], which uses a wavelet-based technique for real-time detection of smoke, the present work proposes the use of the HMT technique, modelling the marginal distribution of the complex wavelet coefficients, to capture local changes in image details, which are assumed to be caused by the presence of steam.

3. Methodology

3.1. Automatic detection of steam in oil sand video images

Fig. 1 shows three different image frames of a reject chute. The images, captured by a video camera, are intended to be used by a computerized system for the automatic evaluation of the ore-size distribution. During the operational hours, the texture pattern of the chute area can change as the oil, dust or small fragments accumulate or attach to the chute. Presence of steam covering the chute also changes the texture pattern. The automatic algorithm for detection of steam in oil sand video images proposed in this paper makes the assumption that the presence of steam covering the reject chute area acts as a blurring process by reducing locally the image details; in other words, steam reduces local extrema values produced by edges and other texture patterns in the wavelet domain. The three different camera views of Fig. 1 were used to illustrate that proposed algorithm for detection of steam is not sensitive to changes in the background image, which in this case are slightly different.

In order to capture the changes in texture caused by the presence of steam in a video frame, the marginal distributions of the wavelet coefficients in small 48×48 labelled regions are modelled by using the HMT technique. The parameters of the derived models are used as input feature vectors to train an SVM classifier, which is then used to detect the presence of steam in the video image. A flowchart is provided in Fig. 2 to better illustrate the proposed algorithm.

3.2. Texture characterization using wavelet transform

3.2.1. Discrete Wavelet Transform

The traditional decimated DWT [12] has been used with great success in many real-world image processing

Please cite this article as: R.J. Ferrari et al., Real-time detection of steam in video images, Pattern Recognition (2006), doi: 10.1016/j.patcog.2006.07.007

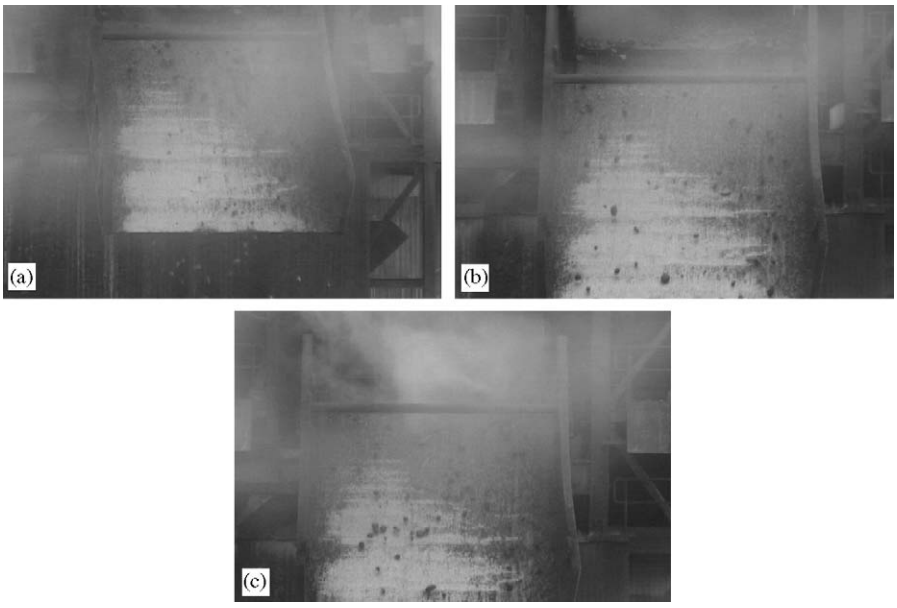

Fig. 1. Video frames of the reject chute during operational hours. (a), (b) and (c) are three different positions of the camera used to demonstrate the results of the proposed algorithm for the detection of steam.

applications for providing multi-scale analysis and sparse representation. In Ref. [13], the author observed that the local maxima of the orthogonal WT of a signal helps detect sharp variations in the signal. Cetin and Ansari [14], proposed an iterative algorithm based on the method of the projections onto convex sets (POCS) for signal recovery from DWT absolute maxima. Due to these properties, in general, only a few large wavelet coefficients containing most of the image energy will be necessary to represent singularities (edges and texture) in an image. A texture pattern, in this case, can be thought of as a region inside the support of a wavelet basis function, which is composed by multiple edges. For natural images, in general, the majority of the coefficients will have small magnitude values, and therefore a histogram plot of the magnitude of the DWT coefficients will have a peak at zero and heavy-tails on the left and right sides [15,16] (see Fig. 3(d) and (e)). Modelling the marginal densities of the magnitude of the wavelet coefficients in two possible states (small or large values) and their persistence or correlation across the scales via statistical HMT models can be seen as a natural extension of the wavelet sub-band energy approach in texture problems.

Despite the success of the DWT in image processing applications, the lack of shift invariance of this transform is a limiting factor in the characterization of texture patterns [17,18]. For the DWT, small spatial changes (rotation or translation) in an image will produce unpredictable wavelet coefficients, making recognition of texture a difficult task since a texture may be presented in the image under any translation. The shift-variance of the DWT also makes the inter-scale persistence property, used by HMT models, no longer hold. Other additional drawbacks of the DWT for the analysis of texture, as demonstrated by Ref. [18], are the poor degree of regularity (smoothness) and symmetry (linear-phase) of the decomposition filters and poor directional selectivity.

Although shift invariance can be obtained by using an undecimated version of the DWT (UDWT) [18], the redundancy imposed by this representation can be prohibitive, especially in real-time image processing applications. Gabor filters have also been used extensively in texture analysis due to their good directional selectivity and optimal lower bound in the joint uncertainty (best trade-off between the spatial and frequency resolutions). However, Gabor wavelet representations (GWRs) are usually highly redundant and the selection of an appropriate bank of filters is highly dependent on the image–frequency content.

3.2.2. *Dual-tree complex wavelets*

The DT-CWT, proposed fairly recently by Kingsbury [19], is a good candidate amongst the DWT, UDWT, and GWR in real-time image application since the DT-CWT provides nearly shift-invariance with a limited redundancy (4:1 for the 2D case—independent of the number of scales) and good directional selectivity (six oriented sub-bands, $\pm 15°$, $\pm 45°$, $\pm 75°$, in comparison to three orientations, $0°$, $45°$, and $90°$, provided by the DWT). Because of the DT-CWT is built by using parallel wavelet trees, this wavelet transform presents excellent degrees of symmetry, which provides linear phase and regularity due to the use of Gabor-like complex type of filters.

The nearly shift-invariant property, valid for the magnitude of the complex wavelets, is obtained by computing parallel wavelet trees with a real bi-orthogonal transform

Please cite this article as: R.J. Ferrari et al., Real-time detection of steam in video images, Pattern Recognition (2006), doi: 10.1016/j.patcog.2006.07.007

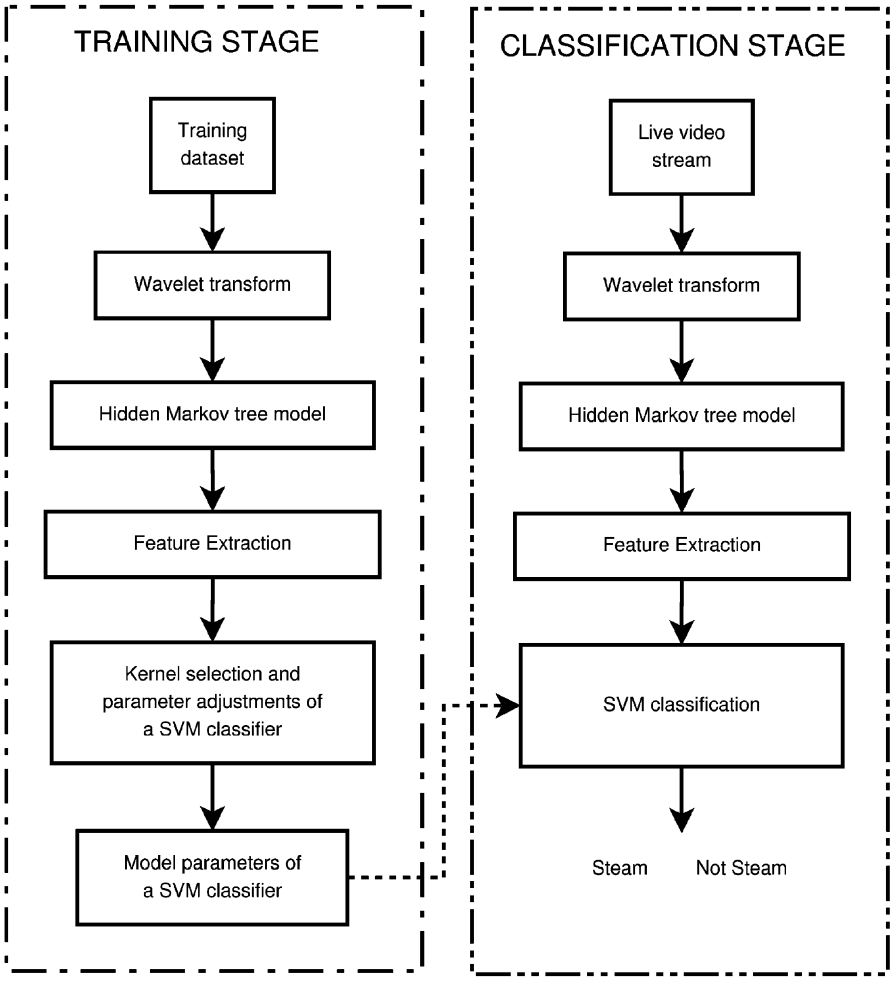

Fig. 2. Flowchart of the technique for the detection of steam in oil sand video images. In the training stage, the parameters of an SVM are determined by using a grid-search procedure.

having double the sampling rate at each scale. Since the complex wavelet transform is approximately shift invariant, the persistence of large/small values of the magnitudes across scales becomes a more valid assumption and more robust HMT models can be built.

3.2.3. Modelling the persistence of wavelet coefficients via HMT Model

The HMT model, applied in the wavelet context, is a statistical model that can be used to capture statistical correlations between the magnitude of the wavelet coefficients across consecutive scales of resolution [15] using the natural quad-tree organization of the DWT.

It is well known from wavelet theory that the probability density function (pdf) of the DWT coefficients can be well modelled by a two zero-mean Gaussian components [15]. One component with small variance, representing a large amount of small coefficient values and a second component with a large variance, representing a small amount of coefficients with large values, as illustrated in Figs. 3(d) and (e). In the case of the DT-CWT, the real and imaginary parts of the wavelet coefficients could also be modelled by two zero-mean Gaussian components since these two parts are approximately in quadrature (i.e., they form an approximate Hilbert pair) and, therefore, their frequency response is almost uncorrelated. However, as shown in Refs. [20,21], a mixture of two Rayleigh pdf's, applied to the magnitude of the complex wavelet coefficients, gives a simpler and more suitable model than a mixture of Gaussians. This fact is also justified by the Gabor-like nature of the DT-CWT filters and due to the non-periodicity of the texture patterns present in the images [22]. Figs. 3(f) and (g) show the Rayleigh components for the DT-CWT case.

In the present work, the marginal distributions of the magnitude of the wavelet coefficients for the DWT and the DT-CWT are modelled, respectively, as a mixture of two-Gaussian and two-Rayleigh components. The persistence property of the wavelet coefficients across scales is formulated by using a first order HMT model. The HMT model is parameterized by the conditional probability stating that

Please cite this article as: R.J. Ferrari et al., Real-time detection of steam in video images, Pattern Recognition (2006), doi: 10.1016/j.patcog.2006.07.007

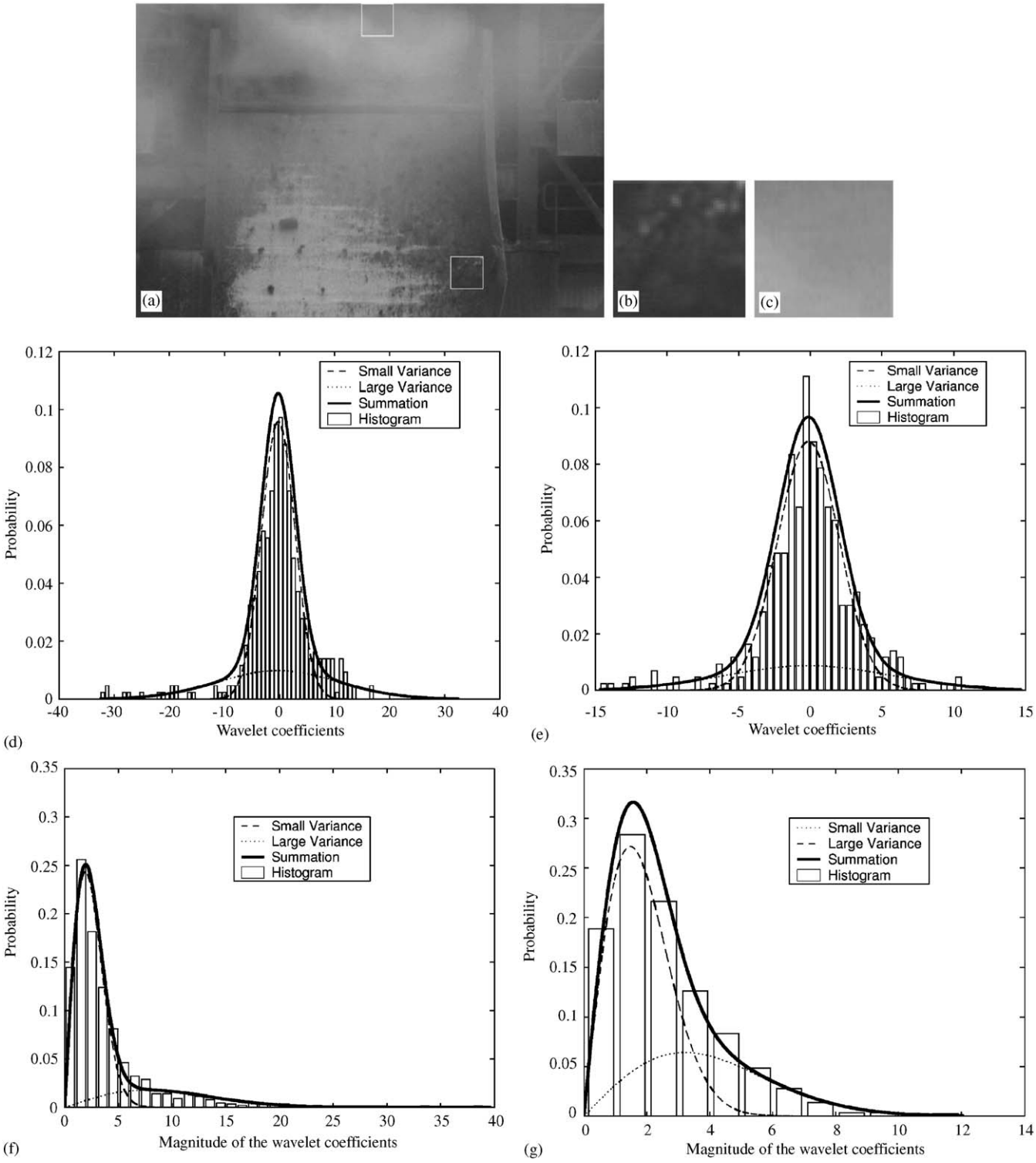

Fig. 3. Modelling of the marginal distribution of the wavelet coefficients via HMT model. (a) Video frame indicating regions samples from (b) the reject chute background and (c) steam, respectively. (d)-(e) Fitting of the marginal distribution of the discrete wavelet coefficients computed for the regions (b) and (c), respectively, by using two-state Gaussian. (f)–(g) Fitting of the marginal distribution of the complex wavelet coefficients computed for the regions (b) and (c), respectively, by using two-state Rayleigh.

the variable S_j is in state m given $S_{\rho(j)}$ is in state n, or, $\varepsilon^{m,n}_{j,\rho(j)} = p(S_j = m | S_{\rho(j)} = n)$ with $m, n = 1, 2$. Indexes j and $\rho(j)$ indicate, respectively, a child and its father node in the quad-tree representation. The state probability of the root J is indicated by $p_{S_J}(m) = p(S_j = m)$ and $\sigma_{j,m}$ are the variances of the mixture components. Due to the admissibility

Please cite this article as: R.J. Ferrari et al., Real-time detection of steam in video images, Pattern Recognition (2006), doi: 10.1016/j.patcog.2006.07.007

condition of the wavelet transform, the mean values of the marginal distributions $\mu_{j,m}$ are zero. For the case of the DT-CWT, the scale parameters of the two-Rayleigh distributions, represented here as $\sigma_{j,m}$ for the sake of consistency in notation, are computed as $\sigma_{j,m} = \sqrt{\sigma^{re}_{j,m} \times \sigma^{im}_{j,m}}$, where $\sigma^{re}_{j,m}$ and $\sigma^{im}_{j,m}$ are the standard deviation of the real and imaginary parts of the complex wavelet transform, respectively.

The parameters, grouped into a vector $\theta = \{p_{S_J}(m), \varepsilon^{m,n}_{j,\rho(j)}, \sigma_{j,m}\}$, are determined by a modified version of the EM algorithm proposed in Ref. [23]. The magnitude of the wavelet coefficients w_i is assumed to follow either one of the two-state Gaussian distributions as

$$f(w_{j,m}|\sigma_{j,m}) = \frac{1}{\sqrt{2\pi\sigma^2_{j,m}}} \exp\left(-\frac{w^2_{j,m}}{2\sigma^2_{j,m}}\right), \quad m = 1, 2 \tag{1}$$

in the case of the DWT or one of the two-state Rayleigh distributions as

$$f(w_{j,m}|\sigma_{j,m}) = \frac{w^2_{j,m}}{\sigma^2_{j,m}} \exp\left(\frac{w^2_{j,m}}{2\sigma^2_{j,m}}\right), \quad m = 1, 2 \tag{2}$$

in the case of the DT-CWT.

The resulting mixture model of the magnitude of the wavelet coefficients for each sub-band is given by

$$P(w_{j,m}) = \sum_{m=1,2} p_{S_J}(m) f(w_{j,m}|\sigma^2_{j,m}). \tag{3}$$

3.3. Steam characterization via statistical modelling of the wavelet coefficients

In the present work, the variance values of the marginal distribution of the wavelet coefficients, showed in Fig. 4, are used as a set of features for the steam classification. Two simplified models were derived from the general hidden Markov framework [15]. In the first model, the possible inter-relationship between the oriented sub-bands is not taken into account; i.e., the wavelet coefficients and the state variables within a particular scale were assumed to be statistically independent. Therefore, the image sub-bands obtained by using either the DWT or DT-CWT were trained independently and hence have their own set of parameters.

In the second model, a further simplification was introduced by tying together all the wavelet coefficients from the wavelet sub-bands for each scale. The tying procedure assumes that the wavelet coefficients for all sub-bands in the same scale arise from the same PDF. By doing that, the number of estimated parameters was reduced to only two for each scale, modelling the two-wavelet states—large and small. In this case, the anisotropic information is dropped in favor of a larger amount of wavelet coefficients to be used in the mixture modelling estimation. The rationale behind this model simplification is the limited spatial resolution

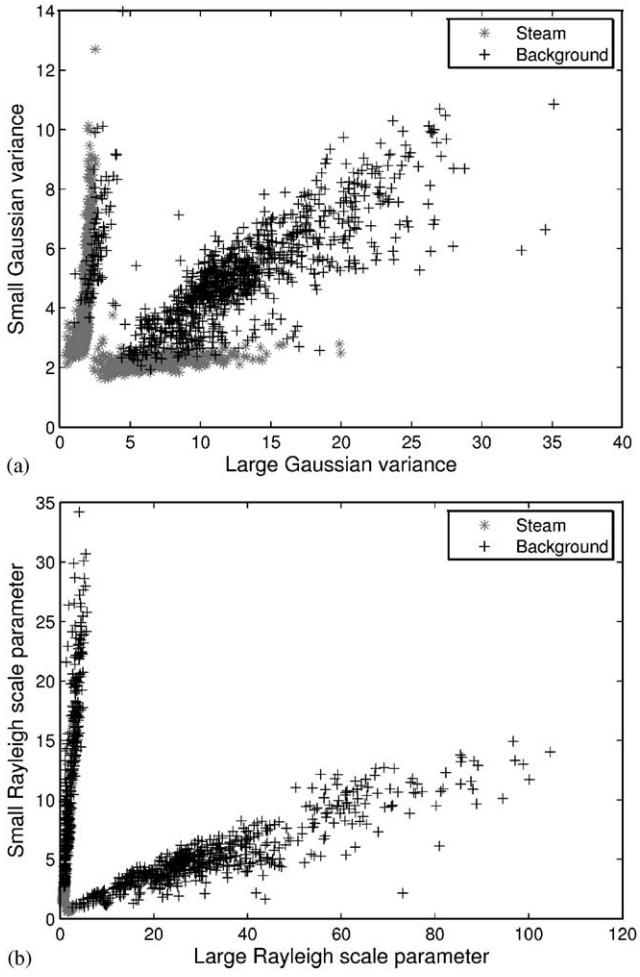

Fig. 4. Scatter plot of the Gaussian and Rayleigh parameters used in the current work as features on the detection of steam and background of the reject chute video images. (a) Large and small variances of the two-Gaussians fitting the marginal distribution of the DWT coefficients. (b) Large and small scale parameters fitting the marginal distribution of the DT-CWT coefficients.

(320 × 240 pixels) in the images used in this work and the need for a reasonable resolution in the detection of a texture. Regions of 48 × 48 pixels were used in the analysis.

Although the parameters of the HMT model are estimated taking into account the persistence of the coefficients using three levels of the wavelet transform, in the present work only the parameters of the second level of decomposition ($\sigma^2_{j=2,m=1}$ and $\sigma^2_{j=2,m=2}$) were used in the steam texture classification. The reason is that in the first level of the decomposition, the wavelet coefficients are mostly related to noise in an image and thus they do not carry relevant information about the texture pattern. For the third level, the use of the parameters has not produced significant improvement in the classification accuracy. This result can be explained by the fact that, in general, natural textures contain a significant amount of energy in midrange frequencies [24].

Detection of steam in the video frames is then performed by using the HMT parameters, estimated for small 48 × 48

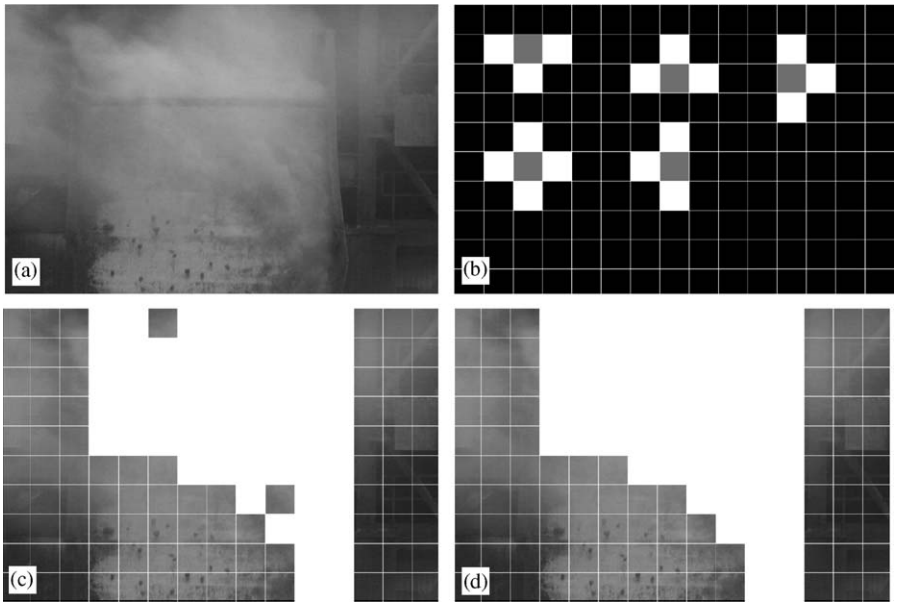

Fig. 5. Spatial constraints used as a post-processing stage to fill holes in the texture pattern. (a) Frame image captured from an oil sand mining site. (b) Disposition of the neighbor regions (in white) evaluated for each "not steam" classified region in the image. (c) Classified image resulting from the SVM classifier. (d) Resulting image after post-processing using information from the classified neighbors. Only the 9×9 centered ROIs were considered in the classification in the reject chute image.

regions, as an input to a supervised classifier. The design of the classifier is described as in the following section.

3.4. Support vector machines

SVMs are supervised machine-learning techniques, recently developed in the framework of statistical learning theory, which have been receiving great attention from the machine-learning community. Due to the robust performance with respect to sparse and noise data, SVMs have been used with success in a large number of applications, including text characterization, face recognition and bioinformatics [25,26].

When used in binary classification problems, SVMs make use of training examples $\mathbf{x}_i$, $i = 1, \ldots, N$, with binary outputs $y_i \in \{+1, -1\}$ in order to learn, in a high-dimensional feature space, a hyperplane $\mathbf{w} \cdot \phi(\mathbf{x}) + b = 0$ which has the maximum margin of separation between the two classes. This is equivalent to solving the following convex-quadratic problem:

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}\|\mathbf{w}\|^2 \\ \text{s.t.} \quad & y_i(\mathbf{w} \cdot \phi(\mathbf{x}_i) + b) \geqslant 1, \end{aligned} \tag{4}$$

where b is an offset parameter, which can be treated as being incorporated into $\mathbf{w}$ by augmenting the feature space vector to $\phi(\mathbf{x}) \to (\phi(\mathbf{x}), 1)$. $\phi(\cdot)$ are called kernel functions which are used to implicitly map the input vectors $\mathbf{x}_i$ into a high-dimensional feature space where the probability of learning a hyperplane with maximum margin of separation is maximized [27]. By using kernel functions, a SVM, which is formulated as a linear discriminant function, can also be applied to non-linear separable problems. In this case, the discriminant function will be determined by the kernel function.

To avoid overfitting the data, in case of noise data, "slack variables" indicated as $\xi_i \geqslant 0$ and a penalty term C are introduced in the hyperplane formulation (Eq. (4)) to relax the margin constraints to

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_i \xi_i \\ \text{s.t.} \quad & y_i(\mathbf{w} \cdot \phi(\mathbf{x}_i) + b) \geqslant 1 - \xi_i. \end{aligned} \tag{5}$$

This new formulation trades off two important goals: (a) to find a hyperplane with maximum margin of separation between classes (i.e., minimizing $\|\mathbf{w}\|$) and (b) to find a hyperplane that separates the data well, or with minimum error, (i.e., minimizing ξ_i). For a practical solution, the problem in Eq. (5) is rewritten by using Lagrange multipliers α_i as

$$\begin{aligned} \min_{w,b} \quad & \frac{1}{2}\sum_{i,j=1}^{m} y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^{m} \alpha_i \\ \text{s.t.} \quad & \sum_{i=1}^{m} y_i \alpha_i, \end{aligned} \tag{6}$$

where $0 \leqslant \alpha_i \leqslant C$ and $K(x_i, x_j)$ are the elements of the Gram matrix K, obtained by evaluating the kernel $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ for all pairs of training inputs. $\alpha_i > 0$ are called

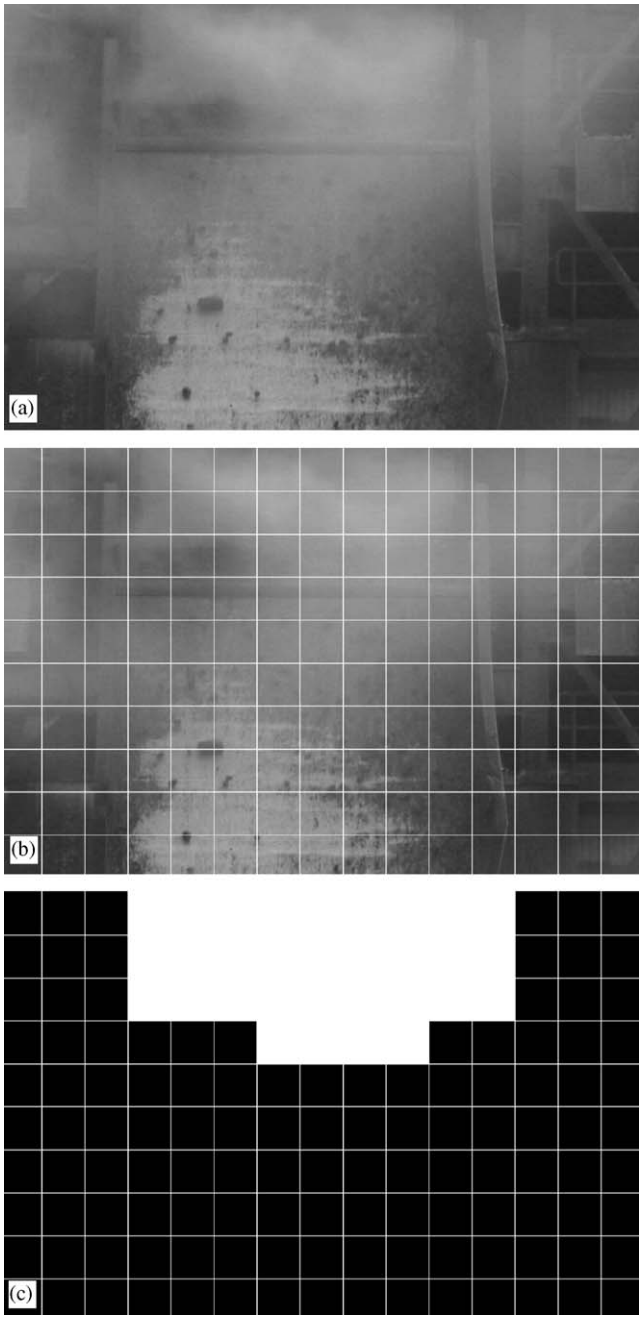

Fig. 6. Example of a frame image with steam used to assess the performance of the algorithm for detection of steam. (a) Original image. (b) Square-grid used to label the image overlayed on the original image. (c) Result of the visual classification.

Table 1
SVM classifiers used in this work and its respective parameters

Kernel	Mathematical expression $K(x_i, x_j)$	Parameters
Linear	$\langle x_i, x_j \rangle$	C
RBF	$e^{-\gamma \|x_i - x_j\|^2}$	C, γ
Polynomial	$(\gamma \langle x_i, x_j \rangle + r)^d$	$C, \gamma, d = 2, r = 0$

support vectors and are the only Lagrange multipliers that contribute in the hyperplane formulation.

Advantageously, the separating maximum-margin hyperplane depends only upon a subset of the original feature vectors known as "support vectors". Therefore, to enhance the performance of a SVM classifier, one should either reduce the set of input features, or find an appropriate kernel and its parameters such that the number of support vectors would be minimized.

3.5. Post-processing using spatial constraints

A reasonable assumption that can be made, regarding the presence of steam in oil sand images, is the spatial-continuity behavior of this texture pattern. In other words, small regions surrounded by opaque steam in an image are also likely to contain steam. By using this simple observation, a simple post-processing algorithm was developed to avoid the presence of unnatural holes in the steam texture, caused by false-negatives from the SVM classification. For the purpose of steam detection, each image frame was subdivided into small regions of 48×48 pixels as illustrated in Fig. 5. Each small region, defined in this paper as a region of interest (ROI), is individually classified as "steam" or "not steam" by using a binary SVM classifier specially tailored for this purpose. Fig. 5(c) illustrates a resulting classified image with two evident holes-false-negative signs. In order to fill the holes, each ROI classified as "not steam" by the SVM classifier is iteratively evaluated and the ROI is reclassified as "steam" if at least three of the four neighbor regions, indicated by the white regions in Fig. 5(b), have been previously classified as steam. Fig. 5(d) shows the image frame in Fig. 5(c) after applying the spatial constraints.

3.6. Image acquisition and database

A recorded 30 min video stream (corresponding to a 54,000 frames) captured by a CCD video camera installed in a Syncrude mine located in Fort McMurray, Alberta, Canada, was used in the present work. The interlaced resolution of the images is 640×480 pixels with 256 gray levels. During the video record time, the camera was moved to three different positions. Video frames were collected with the camera being stationary.

A total of 1000 ROIs (equivalent to 10 frames) for each class—background and steam—were randomly selected from the entire video stream. Only the effective area of the reject chute, defined as the 9×9 centered ROIs, were considered by the classification algorithm (see Fig. 5).

A group of 20 video frames containing steam/non-steam regions (81 ROIs per frame) were manually labelled for the sake of comparison to human performance in detecting the presence of steam. For each frame, the three previous and

Please cite this article as: R.J. Ferrari et al., Real-time detection of steam in video images, Pattern Recognition (2006), doi: 10.1016/j.patcog.2006.07.007

Table 2
The best two average results obtained using a five-fold cross-validation procedure for the linear, RBF and polynomial SVM kernels with features derived from the tied and non-tied DWT/HMT model

Features	Parameter		Average				
	C	γ	Accuracy (%)	#FP	#FN	#SV	TPT (s)
Linear kernel							
Tied	0.01		86.9 ± 0.58	19.6	33	524.2	0.26
Tied	0.1		86.7 ± 0.54	20.6	32.6	497.8	0.74
Non-tied	1		95.6 ± 0.49	5.6	12	187.4	4.72
Non-tied	0.1		96.6 ± 0.68	4.8	12.8	208	0.98
RBF kernel							
Tied	10	1	95.75 ± 0.63	4	13	288.8	2.39
Tied	1	1	95.5 ± 0.16	2.6	15.4	362.8	1.75
Non-tied	10	0.01	97.5 ± 0.38	2.4	7.6	210.2	0.69
Non-tied	10	0.1	95.9 ± 0.18	8.8	7.6	750.4	3.36
Polynomial							
Tied	1	0.01	95.75 ± 0.91	2.6	14.4	177.8	7.29
Tied	1	0.01	95.7 ± 0.54	2.8	14.4	167.6	52.72
Non-tied	10	0.01	97.75 ± 0.34	2.8	6.2	97.2	65.76
Non-tied	1	0.01	97.5 ± 0.63	2.8	7.2	114.2	7.93

FP and FN indicate, respectively, the number of false-positives and false-negatives steam detection. SV is the number of support vectors and TTP stands for the training and processing time measured in seconds.

Table 3
The best two average results obtained using a five-fold cross-validation procedure for the linear, RBF and polynomial SVM kernels with features derived from the tied and non-tied DT-CWT/HMT model

Features	Parameter		Average				
	C	γ	Accuracy (%)	#FP	#FN	#SV	TPT (s)
Linear kernel							
Tied	100		98.5 ± 0.44	2.2	3.8	55.2	4.24
Tied	0.1		98.55 ± 0.36	1.2	4.6	76.8	0.065
Non-tied	0.35		98.65 ± 0.58	1.4	4	58	0.66
Non-tied	8		98.65 ± 0.33	1.8	3.6	49.8	4.33
RBF kernel							
Tied	2	0.25	98.55 ± 0.42	4	3.4	211.8	0.94
Tied	2	8	99.05 ± 0.13	2.6	1	667.4	3.56
Non-tied	2	0.1	98.45 ± 0.67	3	3.2	822.2	4.9
Non-tied	1	0.01	98.1 ± 0.52	2.4	5.2	601.6	2.73
Polynomial							
Tied	1	0.25	98.65 ± 0.33	1.6	3.8	57.2	2.17
Tied	1	0.5	99.55 ± 0.17	2.6	3.2	48.6	10.22
Non-tied	1	0.01	98.15 ± 0.11	1.6	5.8	66.4	87
Non-tied	1	0.05	98.05 ± 0.2	2.4	5.4	55	6

FP and FN indicate, respectively, the number of false-positives and false-negatives signs. SV is the number of support vectors and TTP stands for the training and processing time.

the three posterior frames were also used to give the sense of motion, so that the presence of steam could be better visually assessed by a human. The three authors of this paper assessed together each individual frame on a 14″ computer screen and also on a large projector screen. The steam and non-steam classification was conducted by overlaying a square-grid image on the top of the original image as illustrated in Fig. 6.

4. Experimental results and discussion

4.1. Design of the SVM classifier

Although the optimal selection of the SVM kernel and its respective parameters for a given application is a very active area of research, so far there is no conclusive study. Therefore, in general, the tuning of a SVM is done by using

Please cite this article as: R.J. Ferrari et al., Real-time detection of steam in video images, Pattern Recognition (2006), doi: 10.1016/j.patcog.2006.07.007

Table 4
Average results of the classification of the 20 labelled images (total of 1620 ROIs) using a linear-SVM classifier with features derived from the DT-CWT/HMT tied model with and without using spatial-context constraints

Experiment	Average		
	Accuracy (%)	#FP	#FN
Without spatial-context constraint	89.88 ± 5.17	2.4	5.8
With spatial-context constraint	90.99 ± 4.78	2.65	4.65

either an exhaustive- or a grid-search approach [28]. In both situations, cross-validation is usually applied.

In the present work, three SVM kernels (linear, RBF[1] and second-order polynomial—Table 1) were evaluated by following the methodology proposed in Ref. [28]. The parameter γ of the RBF kernel denotes the width of the Gaussian function. For the polynomial kernel, d, γ and r represent, respectively, the degree, the coefficient and the co-additive constant of the polynomial function. In this work, the order of the polynomial function was set to $d=2$ to avoid overfitting the data by giving high flexibility to the function. The co-additive constant was set to $r=0$ (the default value in the libsvm library [29]). In each case, the regularization parameter C and the parameter γ were first varied using a coarse grid from 2^{-5} to 2^{15} with steps of 2^2. After finding the best grid range, the best SVM kernel parameters were assessed again using a finer grid. For each combination of C and γ parameters in the search grid, the SVM algorithm was considered not to converge and was stopped if the training time exceeds 30 min. A Pentium IV 2.66 GHz, 512 MB machine was used for this task. Mean and standard deviation values for the classification accuracy were obtained by using the held-out test data for a 5-fold cross-validation ([30, Chapter 9, p. 483]).

4.2. Classification results

The cross-validation assessment of the SVM algorithm was performed by using all the ROIs from steam and background regions composing the image database. Table 2 presents the best two average results of the classification obtained for the tied and non-tied DWT/HMT models. Except for the tied model using the linear SVM kernel, the average accuracy for the evaluated kernels is very similar. The non-tied model using the polynomial kernel performed slightly better in both the average accuracy and in the number of support vectors.

The best average results of the classification obtained for the tied and non-tied DT-CWT/HMT models are presented in Table 3. Different from the DWT, the tied models performed slightly better than the non-tied ones in the case of the RBF and polynomial kernels. The number of support vectors were lower in the case of the linear and polynomial kernels.

As can be noticed in Tables 2 and 3, the average accuracy results obtained by using the DT-CWT were always higher than the DWT. The feature plots in Figs. 4(a) and (b) also show that the distribution of the DT-CWT parameters presents lower overlapping between the steam and background classes comparing to the DTW parameters. These results confirm the advantages of using this transform in the characterization of texture pattern as argued in Section 3.2.

A group of 20 frames manually labelled containing steam and non-steam regions (81 ROIs per frame—only the effective region of the reject chute) was used to quantitatively assess the proposed algorithm. The results presented in Table 4 indicate a good agreement (about 90% of correct classification) between the computer classification and the operators opinion. The results also show that, for this group of images, the spatial-context constraint did not improve the average accuracy of the classification.

Figs. 7 and 8 show the results of the steam detection in a sequence of image frames for two different positions of the video camera. For reduction in the computational time, only the effective region of the chute is processed for the steam detection. The area covered by the steam in each frame image can be trivially computed by multiplying the number of ROIs detected as steam regions by the area of each ROI represented by 48×48 pixels in an image.

5. Conclusions

In this paper, we proposed a real-time image processing algorithm for detection of steam in oil sand video images. The algorithm, which works on an individual frame basis, uses the parameters of HMT technique, modelling the marginal distribution of the wavelet coefficients of small square regions of an image, as features to characterize the steam texture pattern. The algorithm was trained and tested using a cross-validation procedure and also compared against a labelled image data set. The results obtained with the proposed algorithm (cross-validation average accuracy of 90% with FP and FN less than 5%) are very promising considering the small number of features used in the classification. For a future study, measures of changes in the chrominance values of the pixels and information about the dynamics of the steam maybe used to improve the current algorithm.

Although, the proposed method performs only binary classification of the steam, the authors intend to extend the method by including an evaluation of the degree of opacity. By determining the steam opacity, a computerized image processing algorithm for ore-size analysis could automatically decide either to recover the original information in the areas covered by steam or simply to discard the

[1] RBF stands for radial basis function.

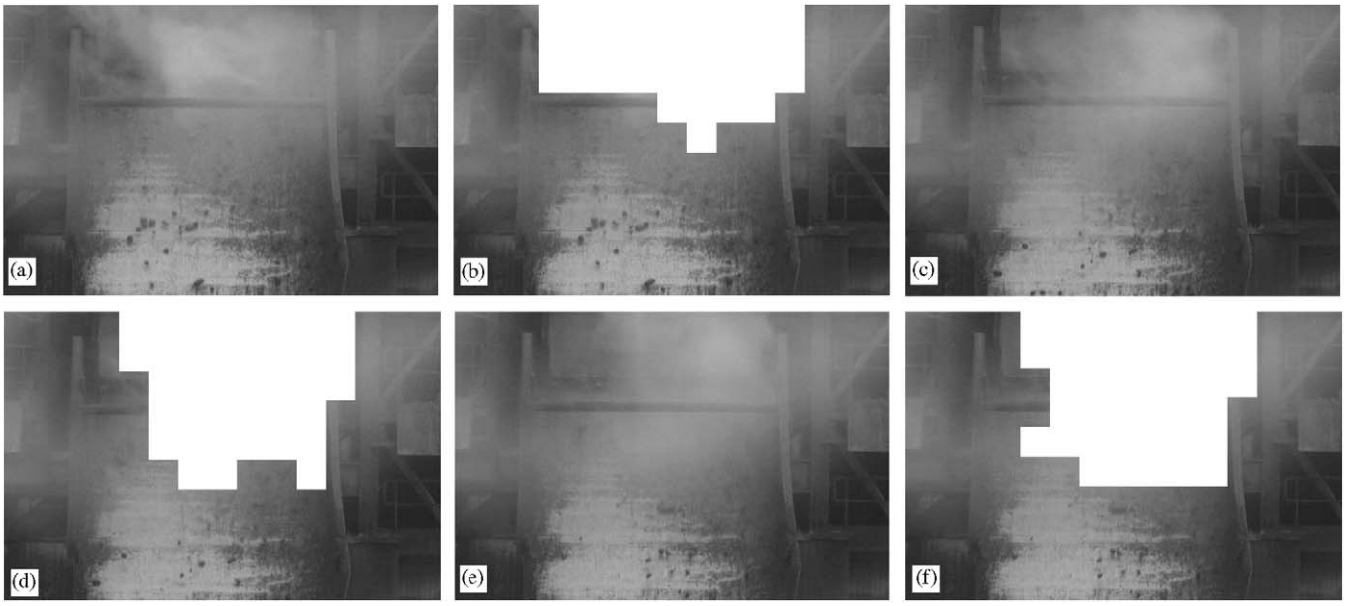

Fig. 7. Sequence of frames from a reject chute in a Syncrude mine—view 1. Left: original images. Right: results of the steam classification using the proposed algorithm.

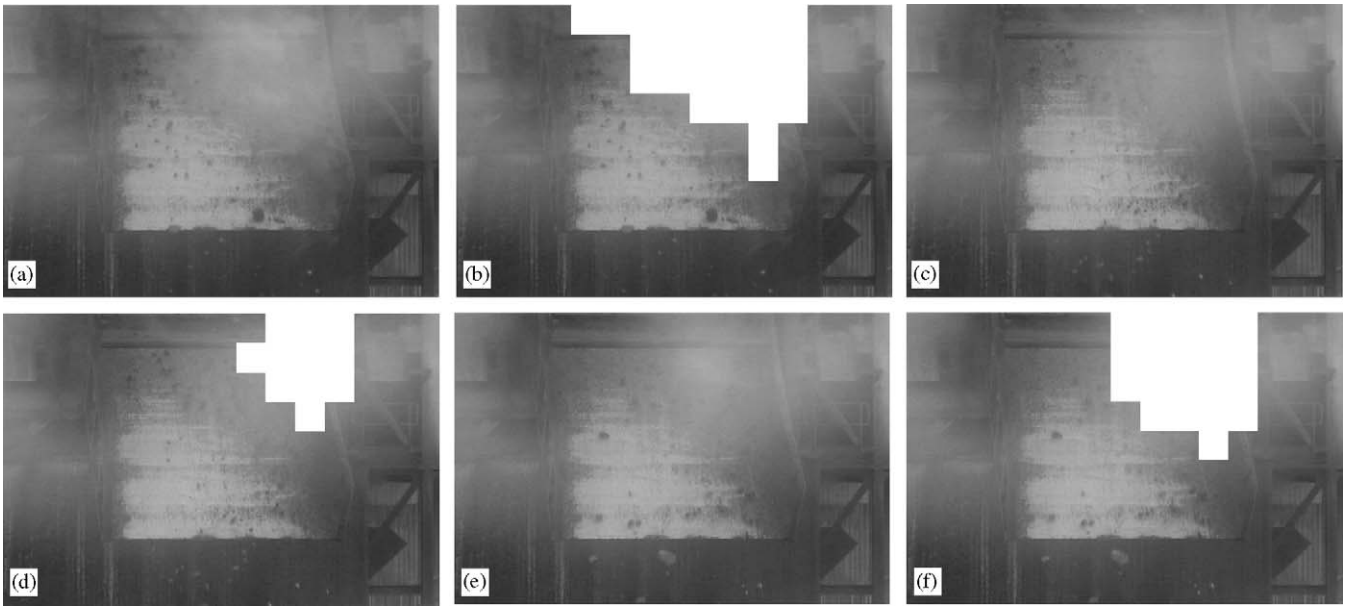

Fig. 8. Sequence of frames from a reject chute in a Syncrude mine—view 2. Left: original images. Right: results of the steam classification using the proposed algorithm.

current video frame during the analysis. A preliminary attempt using probability values provided by an SVM probability mapping method (available in the libsvm library) as an indication of the degree of steam opacity did not produce satisfactory results.

In addition to the importance of detecting steam in oil sand video images, the proposed algorithm may also be adapted to other industrial applications such as fire alarms, cargo smoke detection and other video smoke detection systems.

Acknowledgments

This work was funded by NSERC, iCORE, University of Alberta, Syncrude Canada Ltd., Matrikon Inc. and Centre for Intelligent Mining Systems.

References

[1] F. Dornaika, H. Zhang, Granulometry with mathematical morphology and motion, in: International Conference on Control Science and Engineering, 2003, pp. 18–20.

[2] E. Cabello, M. Sanchez, J. Delgado, A new approach to identify big rocks with applications to the mining industry, Real Time Imaging 8 (1) (2002) 1–9.
[3] F. Yan, H. Zhang, C. Kube, A multistage adaptive thresholding method, Pattern Recognition Letters 26 (2005) 1183–1191.
[4] A. Zadorozny, H. Zhang, M. Jagersand, Granulometry using image transformation techniques, in: 15th International Conference on Vision Interface, Calgary, AB, Canada, 2002, pp. 433–438.
[5] F. Dornaika, H. Zhang, Granulometry using mathematical morphology and motion, in: Proceedings of 2000 IAPR Workshop on Machine Vision Applications, Tokyo, Japan, 2000, pp. 51–54.
[6] T.-H. Chen, P.-H. Wu, Y.-C. Chiou, An early fire-detection method based on image processing, in: Proceedings of IEEE International Conference on Image Processing ICIP-04, vol. 3, 2004, pp. 1707–1710.
[7] B. Töreyin, Y. Dedeoglu, A. Cetin, Wavelet based real-time smoke detection in video, in: Proceedings of the 13th European Signal Processing Conference, EUSIPCO 2005, Antalya, Turkey, 2005, pp. 4–8.
[8] G. Doretto, D. Cremers, P. Favaro, S. Soatto, Dynamic texture segmentation, IEEE International Conference on Computer Vision, vol. 2, Wiley, New York, 2003, pp. 1236–1242.
[9] G. Doretto, A. Chiuso, Y. Wu, S. Soatto, Dynamic textures, Int. J. Comput. Vision 51 (2) (2003) 91–109.
[10] I. Moore, Video Smoke Detection over I.P—Features & Benefits, 19 September 2005.
[11] G. Privalov, SigniFire ⟨http://www.axonx.com/⟩, September 2005.
[12] S. Mallat, S. Zhong, Characterization of signals from multiscale edges, IEEE Trans. Pattern Anal. Mach. Intell. 14 (7) (1992) 710–732.
[13] S. Mallat, Zero-crossing of a wavelet transform, IEEE Trans. Inf. Theory 37 (1991) 1019–1033.
[14] A. Cetin, R. Ansari, Signal recovery from wavelet transform maxima, IEEE Trans. Signal Process. 42 (1) (1994) 194–196.
[15] M. Crouse, R. Nowak, R. Baraniuk, Wavelet-based statistical signal processing using hidden Markov models, IEEE Trans. Signal Process. 46 (4) (1998) 886–902.
[16] J. Romberg, H. Choi, R. Baraniuk, Bayesian tree-structured image modeling using wavelet domain hidden Markov models, IEEE Trans. Image Process. 10 (7) (2001) 1056–1068.
[17] C.-M. Pun, Rotation–invariant texture feature for image retrieval, Comput. Vision Image Understanding 89 (1) (2003) 24–43.
[18] A. Mojsilovic, V. Popovic, D. Rackov, On the selection of an optimal wavelet basis for texture characterization, IEEE Trans. Image Process. 9 (12) (2000) 2043–2050.
[19] N. Kingsbury, Image processing with complex wavelets, Philos. Trans. Roy. Soc. Lon. 357 (1999) 2543–2560.
[20] J. Romberg, H. Choi, R. Baraniuk, N. Kingsbury, A hidden Markov tree model for the complex wavelet transform, Technical Report 2-HIER, 2-REST, Applied and Computational Mathematics—Caltech, (September 2002).
[21] J. Sun, D. Gu, S. Zhang, Y. Chen, Hidden Markov bayesian texture segmentation using complex wavelet transform, in: IEE Proceedings of Vision, Image and Signal Processing, vol. 151, 2004, pp. 215–223.
[22] S. Bhagavathy, J. Tesic, B. Manjunath, On the Rayleigh nature of Gabor filter outputs, in: IEEE International Conference on Image Processing (ICIP), vol. 3, Barcelona, Spain, 2003, pp. 745–748.
[23] J.-B. Durand, P. Goncalves, Statistical inference for hidden Markov tree models and application to wavelet trees, Technical Report 4248, In-stitut National de Recherche en Informatique et en Automatique. ⟨http://www.inria.fr⟩, September 2001.
[24] J. Smith, S.-F. Chang, Transform features for texture classification and discrimination in large image databases, in: Proceedings of IEEE International Conference on Image Processing (ICIP-94), vol. 3, Austin, TX, United States, 1994, pp. 407–411.
[25] C. Burges, A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery, vol. 2, Kluwer Academic Publishers, Boston, 1998, pp. 121–167.
[26] M. Brown, W. Grundy, D. Lin, N. Cristianini, C. Sugnet, T. Furey, M. Ares Jr., D. Haussler, Knowledge-based analysis of microarray gene expression data by using support vector machines, in: Proceedings of the National Academy of Sciences, vol. 97, 2000, pp. 262–267.
[27] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press, Cambridge, 2000.
[28] C. Hsu, C. Chang, C. Lin, A Practical Guide to Support Vector Classification, Guide, Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan, July 2003.
[29] C.-W. Hsu, C.-J. Lin, A comparison of methods for multi-class support vector machines, IEEE Trans. Neural Networks 13 (2) (2002) 415–425.
[30] R. Duda, P. Hart, D. Stork, Pattern Classification, second Ed., Wiley Interscience, New York, 2000.

About the Author— RON KUBE received his B.Sc., M.Sc. and Ph.D. degrees in Computing Science from the University of Alberta. He is currently a Research Associate with Syncrude Canada Limited and an adjunct professor in Computing Science at the University of Alberta. Ron's research interests include novel sensor-based systems and collective robotics.

About the Author—HONG ZHANG received his B.Sc. degree from Northeastern University, Boston, USA, in 1982 and his Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1986, both in Electrical and Computer Engineering. He is currently a Professor in the Department of Computing Science, University of Alberta, and the Director of the Centre for Intelligent Mining Systems. He is the holder of the Syncrude/Matrikon NSERC/iCORE Senior Industrial Research Chair in Intelligent Sensing Systems. His current research interests include robotics, computer vision, image processing, and intelligent systems.

About the Author—RICARDO JOSÉ Ferrari currently works as a medical imaging consultant specialized in the area of breast cancer. He graduated in electrical engineering at the University of São Paulo, Brazil in 1996, and received his M.Sc. in 1998 and his Ph.D. in 2002 from the same University. Between 1999 and 2000, he worked on his Ph.D. project at the University of Calgary, Canada. His studies were in the field of Computer-aided diagnosis for the detection of breast cancer. During his post-doctoral he worked in the detection of MS-lesions in MR images. Ricardo has published more than 26 scientific papers in peer-reviewed conferences and journals, and has served as a reviewer for the many recognized medical imaging journals.

Please cite this article as: R.J. Ferrari et al., Real-time detection of steam in video images, Pattern Recognition (2006), doi: 10.1016/j.patcog.2006.07.007