Wavelet Energy and the Usefulness of its Powers in Motion Detection

Igor VUJOVIĆ, Ivica KUZMANIĆ
University of Split, Faculty of Maritime Studies, Department of Marine Electrical and Information Technologies, Signal Processing, Analysis, and Advanced Diagnostics Research and Educational Laboratory (SPAADREL), Ruđera Boškovića 37, 21000 Split, Croatia
ivujovic@pfst.hr

***Abstract*—The potential for the usage of energy exponents in motion detection from video sequences is explored. The wavelet domain was chosen for the research due to the optimality of Hilbert's space for energy calculations and Parseval's equation for energy equivalence between domains. Five algorithms were considered: wavelet energy motion detection algorithm based on wavelet pairs and buffer, listed in the references, and four which are the contributions of this paper: modification by the application of different wavelet pairs, a modified algorithm without buffer, a modified algorithm without buffer and pairs, newly developed algorithms for energy exponents with and without buffer, but with wavelet pairs. The considered algorithms are background subtraction algorithms modified not to use pixels values, but rather energy/energy exponent backgrounds and the current situation models. These models are described by wavelet descriptors, the introduction of which is the contribution of this paper. They are compared by standard statistical criteria and execution time. The results suggest that an increase in the energy exponent decreases precision, recall and F-measure. However, the percentage of correct classifications remains almost constant. Higher exponentials reduce noise, but are more susceptible to shadows, the waving tree effect and similar abnormalities. Algorithms without buffers are less robust to illumination changes.**

***Index Terms*—discrete wavelet transform, image motion analysis, morphological operations, motion detection, wavelet coefficients.**

I. INTRODUCTION

Energy methods are widely used in signal and image processing (i.e. [1 - 6]). The most popular fields of research are image processing and analysis. Research has shown that the application of energy methods and algorithms in these fields provides simple and efficient algorithms yielding satisfactory results. For example, energy can be used to fuse images [1, 7], detect motion [2], or in biometrics for security purposes [8]. Energy-based measurement of visual quality is presented in [3]. Improvement of edge preservation based on wavelet Ginzburg-Landau energy is covered in [9]. Reference [8] uses wavelet energy for palmprint identification. Wavelet energy is used in fields unrelated to image processing, such as in analysis of stock data [10], power quality signal classification [11], biomedical signals such as snore classification [12] or epilepsy seizure detection [13]. Gait energy is used for human identification in [4]. Wavelet energy is used in damage detection and a system identification approach is presented in [5]. Vibration signal analysis could also be performed by energy methods [6]. Wavelet packets are used to increase the reliability of analysis.

Although a number of papers dealt with the issue of motion detection in video sequences, the powers (exponents) of energy [14 - 18] were not used in any of them. Therefore, this research is original in the sense that the usage of energy exponentials has never been reported before. Our intention is to find out why, by using statistical measures to compare the performance of various energy powers. Since there are no publicly available wavelet motion detection algorithms for energy exponents, we developed and used them to draw conclusions.

The paper is organized as follows. The second section gives the mathematical background. The working hypothesis is set. The third section describes the algorithms used and their modification for purposes of this paper. Differences are explained. The fourth section presents the results. Based on these results, propositions are developed. Finally, conclusions and discussion are given.

II. MATHEMATICAL BACKGROUND

An L_p space may be defined as a space of functions for which the pth power of the absolute value is Lebesgue integral. The p-norm in discrete case can be expressed as (1) [19]:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \tag{1}$$

or in the continuous domain as (2):

$$\|f\|_p = \left(\int_S |f|^p \, d\mu \right)^{1/p} \tag{2}$$

Higher exponents can be expected to produce an effect like an H_∞ norm in systems analysis. This norm extracts the peak of i.e. error signal or transfer function response. It can be defined as (3):

$$\|F(s)\|_\infty = \sup_\omega |F(j\omega)| \tag{3}$$

where supremum can be replaced with maximum in engineering practice. In essence, the supremum should be used, because the maximum could yield an incorrect result, since it provides a solution even when $f \to \infty$. Hardy space (H) has a symbol ∞ due to the fact that maximum amplitude, depending on frequency, can be written as in (4) [10 - 21]:

$$\max_\omega |F(j\omega)| = \lim_{p\to\infty} \left(\int_{-\infty}^{\infty} |F(j\omega)|^p \, d\omega \right)^{1/p} \tag{4}$$

The minimization of the H_∞ norm minimizes errors of the

Digital Object Identifier 10.4316/AECE.2017.02009

frequency component with maximum amplitude. Other components are annulated in power operation.

Why is the calculation of the supremum relevant for motion detection? Because we could hope to maximize a parameter (or minimize an error in i.e. the motion mask), such as precision or percentage of correct classifications or similar. So, why not use it? This will be illustrated on the example of energy differencing in the wavelet domain in the Results section.

The nature of Hilbert space makes it convenient for energy methods. For example, wavelet coefficients should only be squared to obtain an energy measure. In case of motion detection, energy differencing (between current and referent energy representation) is actually least-squares (LS) method, which is optimal in Hilbert space [22]. Fourier transform (FT) is known to be optimal for LS [22]. The wavelet transform (WT) can be interpreted through FT and the same result can be achieved.

Wavelet transform is convenient for energy calculations in image space due to Parseval's relation. This relation can be transformed to DWT (Discrete WT) and expressed with (5) as in [23]:

$$\frac{1}{N}\sum_{t}\left|f(t)\right|^2 = EA + ED \tag{5}$$

EA and *ED* are determined by:

$$EA = \frac{1}{N_J}\sum_{k}\left|a_J(k)\right|^2 \tag{6}$$

$$ED = \sum_{J=1}^{J}\left(\frac{1}{N_J}\sum_{k}\left|d_J(k)\right|^2\right) \tag{7}$$

where: N denotes the sampling frequency, *f(t)* the analysed signal, *a* approximation coefficients, and *d* detailed coefficients. Equation (6) expresses the energy contained in approximation coefficients, while energy contained in the detailed coefficients is expressed by equation (7).

Research in [24] proved that the classical Haar bases are optimal with respect to the ∞-norm of the reconstruction error. The above considerations lead to the conclusion that the H_∞ norm could produce better results. This is formulated as Hypothesis.

Hypothesis 1. Motion detection, as the difference between the reference and current frame or model, generates an error signal. If the error signal is strong, there is a high probability of an object in motion. Weak error signal is indicative of high likelihood of noise. Hence, the H_∞ norm is expected to suppress the noise signal and make the motion signal more segmented and contrasted to the rest of the image/model difference. If this signal is binarized, the result is (8):

$$\left\|Current - Background\right\|_\infty = \begin{cases} 1, motion \\ 0, noise \\ 0, background \end{cases} \tag{8}$$

This hypothesis will be tested in the Results section by increasing the energy powers to establish trends.

III. Motion Detection Energy Algorithms and Modifications for Higher Powers

In order to detect motion in video sequences, a modified background subtraction method was used. All proposed algorithms are background subtraction algorithms. The modification is that, in contrast to the classic background method, no pixel values were used, but energy or energy exponent values of the wavelet approximation coefficients that could be linked with the corresponding pixels.

Video sequences are in RGB (red-green-blue) colour space, which is the space of further work without transformation of other colour spaces. Hence, all further formulas are applied to all 3 colour components separately, with the exception of the identification of the maximum, which is the same for all 3 colour-components. All matrixes are thus actually realized in 3D in the Matlab environment. The modified algorithm has similarities with [25], but there are some simplifications in the sense of less morphology and thresholds, there are no wavelet pairs, but only one wavelet, and there are no energy buffers. Furthermore, it is compared with the same algorithm as in [25], and with the modified [25] in order to use powers of energy and compare the results.

Algorithms considered are (in summary):

- Algorithm 1: the wavelet energy motion algorithm from [25], and modification with inclusion of lazy-haar, db2-db8 combinations instead of only db2-db7 (wavelet designations are taken from Matlab),
- Algorithm 2: a new algorithm which does not include wavelet pairs, but single wavelet, does not have energy buffer, but does use energy and mean square error (MSE) method for motion detection,
- Algorithm 3: a modified algorithm 1, which includes pairs, but does not include an energy buffer,
- Algorithm 4: a new algorithm, which includes wavelet pairs, does not include an energy buffer, and incorporates energy exponents. Motion detection is executed by the modification of the MSE to exponents' differences, and
- Algorithm 5: a new algorithm, which includes wavelet pairs, energy exponent buffer, and energy exponent differences for motion detection.

Test algorithms have common basics and differences, and can be described as follows:

Step 1: Initialization and background energy calculation are included in all five algorithms.

Background energy is calculated as the average energy of the first 20 frames in a video sequence (the algorithm from [25] can use any number). The energy of a frame is obtained by a two-dimensional lifting wavelet transform (2D-LWT) and the square of the wavelet approximation coefficients. Finally, the obtained energy is normalized. This process is performed at the second level of wavelet decomposition. The procedure at the second level of decomposition can be expressed with several equations. The first is (9):

$$ap\_coef = \sum_{i=1}^{length(bg\_ref)} ap_i \tag{9}$$

where *ap_coef* is the sum of approximation coefficient matrixes with the quarter of the size of the original frame, and ap_i the matrix of the approximation coefficients of the i-th frame in the training sequence. The expression *length (bg_ref)* denotes the length of the training sequence, which can vary depending of the availability of the non-dynamic part of the sequence, and time available for training. The

background energy model is defined by (10):

$$E_{bg} = \frac{ap\_coef^2}{length(bg\_ref)} \tag{10}$$

where E_{bg} is the matrix of background energy at spatial positions with the quarter of the resolution of the original frame. This matrix is normalized by the maximum matrix member (11):

$$E_{bg\_norm} = \frac{E_{bg}}{\max(E_{bg})} \tag{11}$$

The procedure from (9) to (11) is used in algorithm 2, developed specifically for this paper. Other algorithms use ap_coef1 and ap_coef2 for 2 wavelets in a pair. Total ap_coef^2 (from (10)) is the sum of energies of wavelets 1 and 2.

The result of (11) is called the wavelet energy descriptor of the background (WEDB). This descriptor could be used in other algorithms.

Step 2: The calculation of the energy of the current frame is used in algorithms 2, 3, and 4. Algorithm 1 uses the current energy buffer, which is the average energy of L frames, including the current frame and $L - 1$ previous frames. This means that instead of ap_coef_current in (12), it uses average energy. Algorithm 5 uses a current energy exponent buffer, which is the average energy exponent of L frames, including the current frame and $L - 1$ previous frames instead of ap_coef_current in (12).

The calculation is performed by equation (12), and the obtained result further normalized by (13):

$$E_{current} = (ap\_coef\_current)^2 \tag{12}$$

$$E_{current\_norm} = \frac{E_{current}}{\max(E_{current})} \tag{13}$$

Thresholding of $E_{current\_norm}$ with minimaxi threshold selection scaled with 1/4π. The minimaxi selection uses a fixed threshold chosen to yield minimaxi performance for mean square error against an ideal performance. The minimaxi estimator realizes the mean square error (MSE) obtained for the worst function in a given set. By this criterion, T_1 is obtained. This threshold is scaled to obtain the final threshold (14):

$$T = \frac{T_1}{4\pi} \tag{14}$$

and current normalized energy is expressed with (15):

$$E_{current\_norm} = \begin{cases} E_{current\_norm}, E_{current\_norm} \geq T \\ 0, E_{current\_norm} < T \end{cases} \tag{15}$$

The thresholding operation is performed in Matlab using the soft method. The energy descriptor of the current scene situation is expressed by (15) and is called a wavelet current energy descriptor (WCED).

Step 3: Motion detection by energy difference is performed in algorithms 2 and 3. Algorithm 1 uses the difference of current buffer energy and background energy to obtain E_{differ} in (16). Algorithms 4 and 5 use the difference of energy exponents to obtain E_{differ} in (16). Furthermore, algorithm 5 makes use of the current buffer of energy exponents.

Energy differencing can be expressed with (16):

$$E_{differ} = \left| E_{current\_norm} - E_{bg\_norm} \right| \tag{16}$$

This step differs from [25], because there is no energy buffer. The energy buffer contains the average energy of several consecutive frames and has the function of the current frame in equivalent step as (16). Buffer has an advantage in illumination variation suppression, but detects motion with a small time delay. The length of the buffer is denoted with L in the Results section. $L = 0$ means that there is no buffer. On the other hand, if i.e. $L = 3$, it means that the current and the two preceding frames are used to calculate buffer energy.

The result of (16) is thresholded similar to $E_{current\_norm}$. The procedure results in the obtainment of optimal minimaxi threshold T_2. In order to investigate the effect of this threshold on the final results, the threshold is scaled by different numbers:

$$T_d = \frac{T_2}{thr \cdot \pi} \tag{17}$$

where *thr* varies from 1 to 8 with 0.05 step. In the Results section, the best results are obtained by different *thr*, which is discussed further. The thresholding yields (18):

$$E_{differ} = \begin{cases} E_{differ}, E_{differ} \geq T_d \\ 0, E_{differ} < T_d \end{cases} \tag{18}$$

where E_{differ} represents motion (sometimes referred to as motion energy).

Step 4: Binarization of the energy difference (equivalent of motion mask) is performed in algorithms 1, 2 and 3. Algorithms 4 and 5 use the binarization of the energy exponent difference.

Binarization is expressed by (19):

$$E_{differ\_mask} = \begin{cases} 1, E_{differ} > 0 \\ 0, E_{differ} = 0 \end{cases} \tag{19}$$

where E_{differ_mask} represents the motion mask (equivalent, not in the classic sense).

Note that in (10) the actual nature of the noise is unknown, because we are in the process of obtaining the reference background energy model. Furthermore, averaging is also a method of noise-removal. On the contrary, denoising is used in (15), because there is only one current frame.

Step 5: Morphological filtering is performed by all considered algorithms.

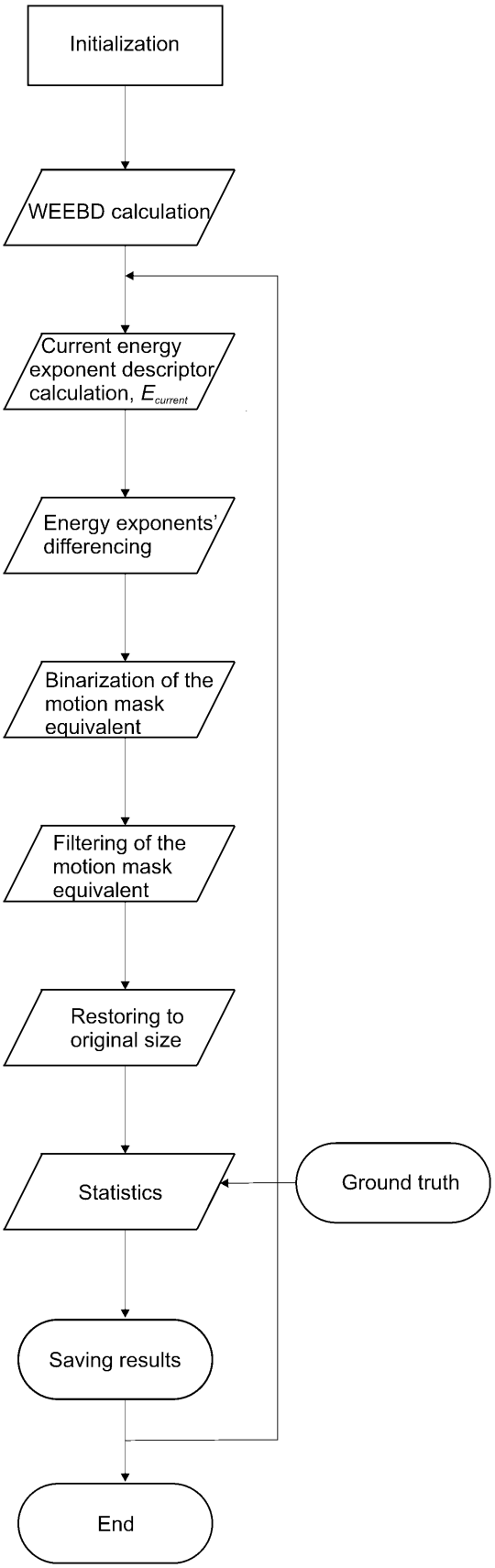
Morphological filtering is performed with the following expression (20):

$$E_{m\_open} = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \left\langle E_{differ\_mask}, MO \right\rangle \tag{20}$$

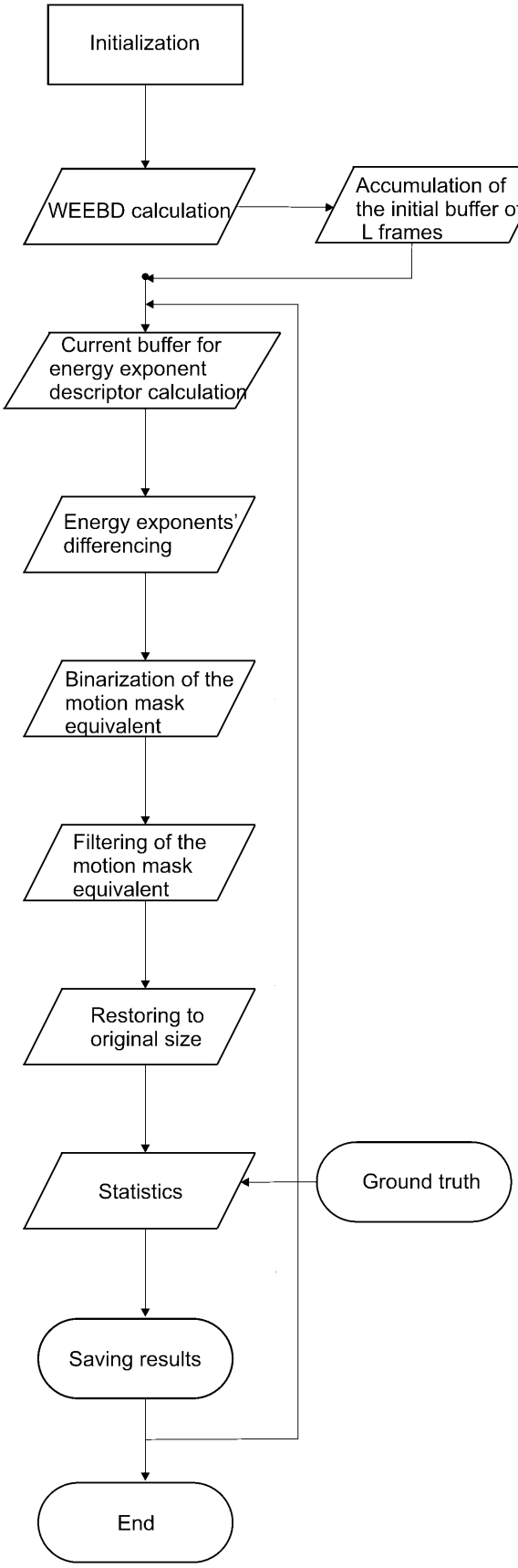
where E_{differ_mask} represents a binarized motion mask matrix (E_{differ}), and *MO* represents the morphological opening kernel. The same procedure is performed for all three colour channels. *MO* operation is used to suppress any artefact pixels caused by noise in a binarized motion mask matrix.

Step 6: The motion mask is transformed to original size by dyadic upsampling and morphological dilatation. These operations are repeated twice to obtain the original size. The dyadic upsampling is performed by a mixed procedure, in which both rows and columns are upsampled. This step is

performed by all the considered algorithms instead of classic inverse LWT.



(a)



(b)

Figure 1. General workflow of the developed algorithm for energy exponents: (a) Algorithm 4, (b) Algorithm 5

Step 7: Statistics are calculated for all thresholds and put into vectors, facilitating access to the result for a specific threshold. The statistics parameters are obtained by comparison with the ground truth results available in the

public dataset (IEEE Workshop on Change Detection – dataset for 2012 Conference, available at: http://wordpress-jodoin.dmi.usherb.ca/dataset2012/). This step is performed in all considered algorithms.

The general workflow of Algorithm 4 is presented in Fig. 1(a), and Algorithm 5 in Fig.1(b).

Algorithm variations are performed in step 1, equation (10), and step 2, equation (12), where square is replaced with other exponents (from 3 to 12), or generally, the procedure performed can be described with (21) and (22):

$$E_{bg} = \frac{ap\_coef^n}{length(bg\_ref)} \tag{21}$$

$$E_{current} = (ap\_coef\_current)^n \tag{22}$$

where n =2, 3, ...

The expression (21) is called wavelet energy exponent background descriptor (WEEBD). The expression (22) is called wavelet current energy exponent descriptor (WCEED).

An energy buffer can be used in the general scheme of our algorithm. It is easily programmable with small differences. Differences between the algorithm with and without accumulator are:

- In step 3, instead of (16), we have:
$$E_{differ} = \left| E_{accumulator} - E_{bg\_norm} \right| \tag{23}$$
- Buffer should be initiated before step 2.
- New step 2 is similar to original step 2, with the difference that current energy is added to the accumulator, after which new average accumulator value is obtained, which is used instead of the current energy for differencing in foreground segmentation step 3.

IV. Results

All experiments were performed on a Dell notebook with Intel® Core™ i5-4200U CPU @1.6 GHz 2.3 GHz, RAM 4 GB, 64 bit Windows 7 Professional.

Considering the sheer number of wavelet families, it should be noted that, due to results in [24], the Haar wavelet was used for the preliminary experiments. Wavelet choices are expanded later in the research process to include db2-db8 (combination of execution speed and edge detection with more vanishing moments), and haar-lazy (for execution speed) pairs.

For comparison purposes, algorithm from Section 3 and the corresponding modified algorithms are compared with statistical measures [14], [25 - 26] expressed with equations (24 - 28):

$$PCC = \frac{TP + TN}{TP + TN + FP + FN} \tag{24}$$

$$precision = \frac{TP}{TP + FP} \tag{25}$$

$$recall = \frac{TP}{TP + FN} \tag{26}$$

$$F = \frac{2 \times precision \times recall}{precision + recall} \tag{27}$$

$$F = \frac{2TP}{2TP + FN + FP} \tag{28}$$

where *PCC* is the percentage of correct classifications, *TP* true positives, *TN* true negatives, *FP* false positives, *FN* false negatives.

The large quantity of data obtained in the research prevents us from presenting them all. Therefore Tables III to VI show only top 10 results.

Results of the preliminary research with the developed algorithm 2 are shown in Tables I and II.

Results in Table III are obtained by algorithms 1, 3, 4, and 5.

Results in Table IV are obtained by algorithms 1, 4, and 5. Results in Table V are obtained by algorithms 1 and 5. Results in Table VI are obtained by algorithms 1 and 5.

Table I presents experimental results for the dataset (IEEE Workshop on Change Detection – dataset for 2012 Conference, available at: http://wordpress-jodoin.dmi.usherb.ca/dataset2012/), sequence "Park". Wavelet energy (square of wavelet coefficients) is denoted with *E* in the table.

PCC can be seen to remain almost constant, no matter what energy powers are used. This can be a bit confusing. If only PCC is observed, one could arrive at the wrong conclusion – that nothing is happening. However, other statistical measures (precision, recall and F-measure) can be seen to deteriorate for higher powers. This can only be explained if TP decreases (which is directly proportionate to precision, recall and F), while TN increases.

Based on observations from Table I, we arrive at conclusions summarised as Proposition 1 and 2.

Proposition 1. When using energy powers, *PCC* remains constant.

Proposition 2. The use of energy powers causes an increase in *TN* and a decrease in *TP*.

Remark 1. Proposition 2 is math, because precision, recall and *F* proportionally depend on *TP*. *PCC* depends on *TP* + *TN*, which is constant. If *TP* is smaller, and the sum is the same, *TN* must increase.

Observation 1. Maximum precision, recall, and F-measure decrease with the increase of power.

The question is why does *TP* decrease with the increase of powers. It might be due to the increase of the effective threshold or increase of noise. Should we adapt the threshold to increase *TP*? Or, is this irrelevant, since the threshold corresponds to the value of the differencing matrix members?

To further investigate relationships between energy powers, we used the absolute difference of the wavelet coefficients, which can be expressed as $\sqrt{E}$ (equation (4) is changed so that power is p = 1 instead of 2). Although the result for maximum *PCC* is improved by mere 0.03%, it is an improvement nevertheless. Precision is 8.97% higher than when wavelet energy is used. The recall is improved by 7.86%. A great surprise is the improvement of the *F*-measure by 41.72%. Execution speed is increased by 3.55% to 24.85 *FPS* (frames per second). It seems that absolute difference yields better results than energy differencing or any other power.

TABLE I. RESULTS OF THE EXPERIMENT: RELATIVE CHANGE IN STATISTICAL MEASURES (HAAR WAVELET, REFERENT ENERGY IS 100%, "TRAFFIC" SEQUENCE; MAXIMUM VALUE, ALGORITHM 2)

	PCC	Precision	Recall	F	FPS
E	100	100	100	100	100
$\sqrt{E^3}$	99.9795	75.8791	50.1282	51.5679	95.4699
E^2	99.0712	62.1838	19.1321	35.8885	95.2491
$\sqrt{E^5}$	99.9795	45.2807	7.6956	21.4286	92.5967
E^3	99.9795	41.8569	3.8478	12.5436	95.2016
$\sqrt{E^7}$	99.9795	40.6231	2.3087	9.2334	97.0526
E^4	99.9795	27.7298	1.5605	7.4042	95.0066
$\sqrt{E^9}$	99.9795	23.5040	0.9833	5.3136	96.5913
E^5	99.9795	21.8384	0.7482	4.8780	97.4497
E^6	99.9795	19.7100	0.4061	3.0488	97.9177

Maximum PCC is obtained for various thresholds, as seen in Table II. A further inconvenience is that the maximums of various statistical measures are not obtained for the same thresholds. This means that a common criterion for maximum performance cannot be established, but only an average threshold.

TABLE II. THRESHOLDS FOR MAXIMUM STATISTICAL PARAMETERS ("TRAFFIC" SEQUENCE, HAAR WAVELET; FOR THE CORRESPONDING MAXIMUM VALUE, ALGORITHM 2)

	PCC	Precision	Recall	F	FPS
E	3	3	8	3.7	1.2
$\sqrt{E^3}$	1.75	2.85	8	8	8
E^2	2.2	3.05	8	8	5
$\sqrt{E^5}$	2.55	3.45	8	8	3.1
E^3	3.05	3.1	7.7 - 8	7.7	3.05
$\sqrt{E^7}$	3	3.65	8	8	6.35
E^4	3.05 - 3.2	3.8–3.9	7.8 – 8	7.65	3.25
$\sqrt{E^9}$	3.45 - 3.5	4.4	8	7.95	1.3
E^5	3.95	4.35	7.9 – 8	7.75	6.45
E^6	3.95 - 4.5	5.45–5.85	7.25 – 8	7.25–7.3	6.3

An increase in energy power means wider range thresholds and worse statistical criteria. For E^6, all statistical parameters have the range of thresholds. Ranges have intersection for F and recall. Other measures do not have intersections in threshold ranges.

After these preliminary experiments, additional experiments were performed using two wavelet pairs instead of one and buffer variations, and an experiment in which no buffer was used. Finally, an additional sequence is used (sequence "PeopleInShade", IEEE Workshop on Change Detection – dataset for 2014 Conference, available at: http://wordpress-jodoin.dmi.usherb.ca/dataset2014/).

Results for sequence "PeopleInShade" are shown in Tables from III to VI. There represents results obtained by modifications in the algorithm from [25].

TABLE III. TOP 10 PARAMETER COMBINATIONS WITH REGARD TO FPS FOR SEQUENCE "PEOPLEINSHADE" (ALL TOP-TEN ARE IN HAAR-LAZY COMBINATION OF WAVELETS; MAXIMUM VALUE, ALGORITHMS 1, 3, 4, 5)

L	Energy	PCC	Precision	Recall	F	FPS	Rank FPS
4	E	0.9236	0.336	0.5771	0.1428	45.40	1
3	E	0.9236	0.342	0.5753	0.1424	43.839	2
6	E	0.9236	0.3402	0.5799	0.1433	42.35	3
5	E	0.9234	0.3328	0.5786	0.143	42.24	4
0	E	0.9262	0.6181	0.5762	0.1672	38.803	5
0	$\sqrt{E^5}$	0.9221	0.3333	0.0958	0.044	37.96	6
3	E^4	0.9196	0.1864	0.0399	0.0269	37.88	7
3	E^5	0.9199	0.2161	0.0585	0.0335	37.82	8
0	$\sqrt{E^7}$	0.9215	0.2895	0.0477	0.0347	37.78	9
0	E^2	0.9226	0.4053	0.1585	0.0614	37.77	10

TABLE IV. TOP 10 COMBINATIONS OF PARAMETERS WITH REGARD TO PCC FOR SEQUENCE "PEOPLEINSHADE" (ALGORITHMS 1, 3, 4, 5)

L	Wavelets	Energy	PCC	Rank PCC
0	haar-lazy	E	0.9262	1
3	haar-lazy	E	0.9236	2
4	haar-lazy	E	0.9236	2
6	haar-lazy	E	0.9236	2
4	haar-lazy	$\sqrt{E^3}$	0.9235	5
5	db2-db8	E	0.9235	5
5	haar-lazy	E	0.9234	7
0	haar-lazy	E^2	0.9226	8
0	haar-lazy	$\sqrt{E^3}$	0.9224	9
3	haar-lazy	$\sqrt{E^3}$	0.9224	9

It can be seen that the most efficient is the original algorithm form [25], and modifications are at 6th place in Table III, 5th place in Table IV, 7th place in Table V, 1st place in Table VI.

TABLE V. TOP 10 COMBINATIONS OF PARAMETERS WITH REGARD TO RECALL AND F-MEASURE FOR SEQUENCE "PEOPLEINSHADE" (ALGORITHMS 1 AND 5)

L	Energy	Wavelets	Recall	F	Rank recall	Rank F
6	E	haar-lazy	0.5799	0.1433	1	2
5	E	haar-lazy	0.5786	0.143	2	3
4	E	haar-lazy	0.5771	0.1428	3	4
0	E	haar-lazy	0.5762	0.1672	4	1
3	E	haar-lazy	0.5753	0.1424	5	5
5	E	db2-db8	0.5365	0.1375	6	6
4	$\sqrt{E^3}$	haar-lazy	0.3133	0.104	7	7
3	$\sqrt{E^3}$	haar-lazy	0.308	0.1012	8	8
6	$\sqrt{E^3}$	haar-lazy	0.3073	0.1011	9	9
5	$\sqrt{E^3}$	haar-lazy	0.3071	0.101	10	10

Best speed is obtained with L = 4 in haar-lazy pair and with energy without powers (square of coefficients), namely 45.4 fps in setup configuration in Matlab.

A higher F-measure is obtained without buffer, also in the haar-lazy, and with energy without powers, namely 0.6181.

The highest recall is obtained by L = 6, haar-lazy, energy without powers, namely 0.5799.

The highest precision is obtained with the db2-db8 combination, L = 5, and $E^{4.5}$, which is 0.914. The runner-up is haar-lazy without buffer, which is 0.6181, and energy without powers.

PCC is the best with the same parameters, namely 0.9262.

TABLE VI. TOP 10 PARAMETER COMBINATIONS WITH REGARD TO PRECISION FOR SEQUENCE "PEOPLEINSHADE" (ALGORITHMS 1 AND 5)

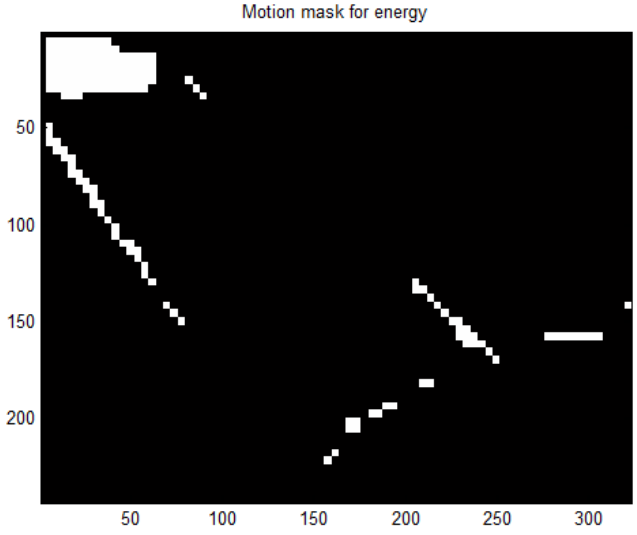
L	Energy	Wavelets	Precision	Rank precision
5	$\sqrt{E^9}$	db2-db8	0.914	1
0	E	haar-lazy	0.6181	2
4	$\sqrt{E^3}$	haar-lazy	0.4336	3
0	E^2	haar-lazy	0.4053	4
3	E	haar-lazy	0.342	5
6	E	haar-lazy	0.3402	6
4	E	haar-lazy	0.336	7
0	$\sqrt{E^5}$	haar-lazy	0.3333	8
5	E	haar-lazy	0.3328	9
5	E	db2-db8	0.3308	10

Fig. 2 shows the comparison of various energy powers through the motion mask for the same threshold and frame 1120. It is only a random example of algorithm operation.

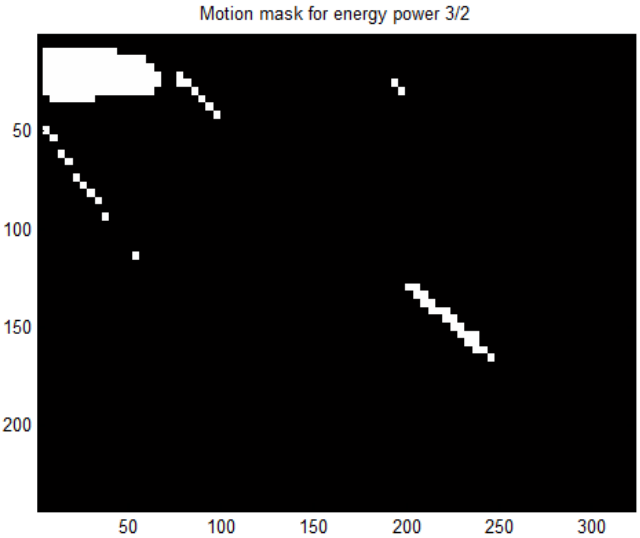
As seen in Fig. 2, shadows and wave-tree effect are

supressed with the increase in powers, as is foreground detection. Therefore, an increase in power is in itself useless for precise motion detection, but can be used in combination to reduce a part of the negative phenomena occurring in outdoor traffic video surveillance. This is merely an observation, not mathematical proof, and must be taken with reservations. Further evaluation in this area could be performed in new researches.

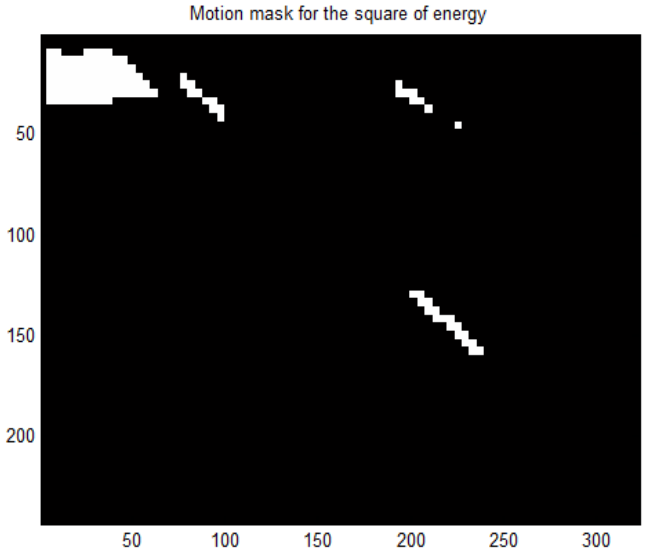
(a)



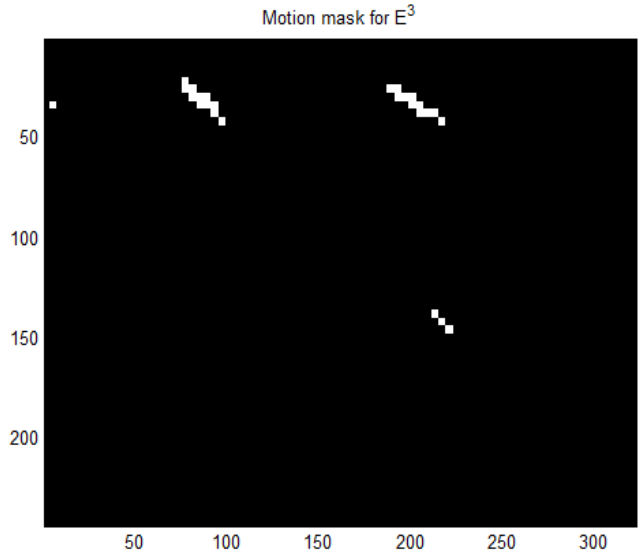
(b)



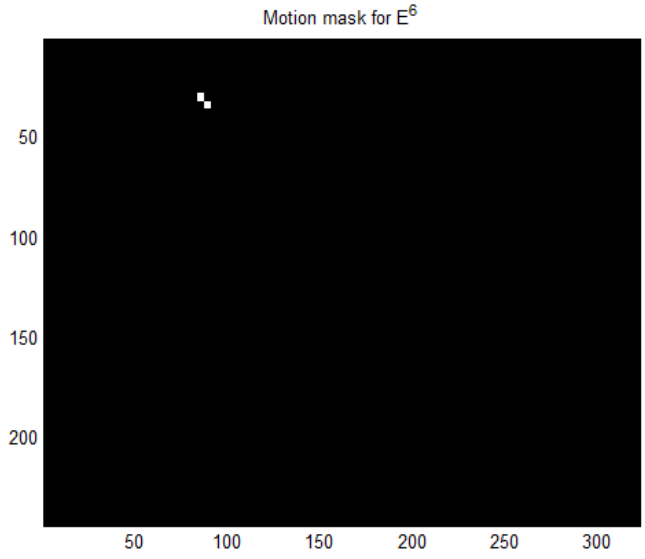
(c)



(d)



(e)



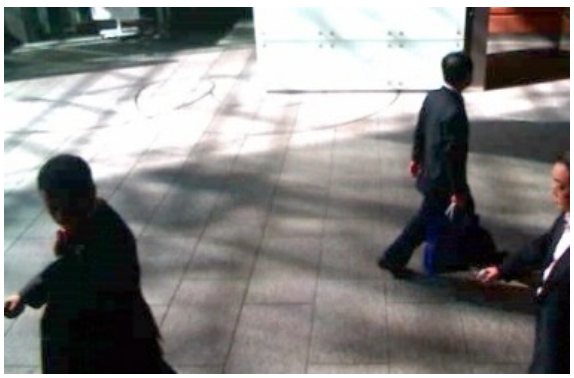
(f)

(g)

Figure 2. Comparison of various powers for "Traffic" sequence: (a) the original frame; (b) energy power = 1 (normal energy); (c) power = 3/2; (d) power = 2; (e) power = 3; (f) energy power = 6; (g) ground truth

Three frames and the ground truth for two frames with motion are shown in Fig. 3 to illustrate the complexity of sequence "PeopleInShade". It is given in order to access the histogram, contrast, etc. Objects in motions are people. The scene is problematic owing to the number of shadows appearing and disappearing during the sequence.

(a)



(b)

(c)

(d)

(e)

Figure 3. "People in Shade" sequence: (a) the first frame; (b) frame 333; (c) ground truth for the frame 333; (d) frame 1110; (e) ground truth for the frame 1110 (video sequence is publicly available at: http://wordpress-jodoin.dmi.usherb.ca/dataset2012/)

Table VII shows examples of the execution times for two wavelets through all powers examined. It is measured with Matlab tic-toc function. Considering Windows time measurement problems, these results can obviously be considered merely framework information, rather than exact measurements. However, algorithms researched can be concluded to be capable of use in real-time.

TABLE VII. REAL FPS PERFORMANCE (TIME IS RECIPROCAL FROM FPS) FOR THE SEQUENCE "PEOPLE IN SHADE"

Power	**FPS for Haar**	**FPS for db2**
E	46.0410	40.6795
$\sqrt{E^3}$	38.5919	37.6956
E^2	41.7577	37.6187
$\sqrt{E^5}$	42.0883	37.6564
E^3	46.7987	37.5902
$\sqrt{E^7}$	46.8941	37.6094
E^4	46.7534	38.3838
$\sqrt{E^9}$	46.8692	38.0920
E^5	46.9659	37.7700
E^6	46.6568	37.9582

Finally, we will shortly discuss the dependence of motion detection accuracy on illumination variations. This is a serious issue addressed by many authors [16, 27-30]. Fig. 4 shows the results of experimental measurement of illumination variations. The experiment was performed under indirect sunlight on a sunny day. The scene is static without motions. Hence, the average pixel value should be constant. However, the results show variations in the illumination of the natural light source.

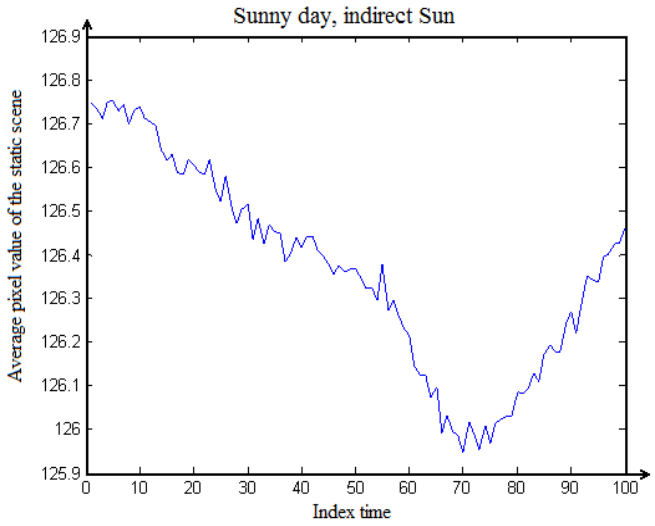
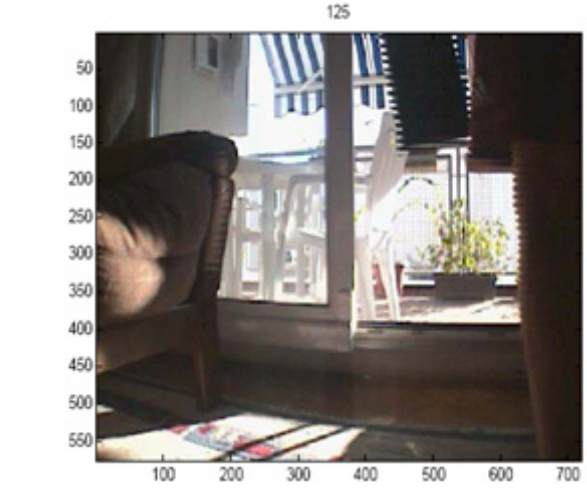


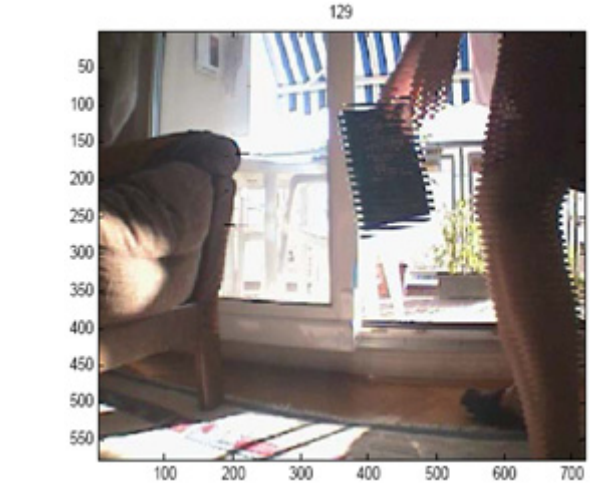
Figure 4. Experimental measurement of illumination variations

Such changes in illumination could intuitively be expected to affect frame differencing and background subtraction algorithms. Namely, a change in pixel value between frames is different from zero. In case of slight illumination changes, the changes in pixel value might not exceed the motion detection threshold.

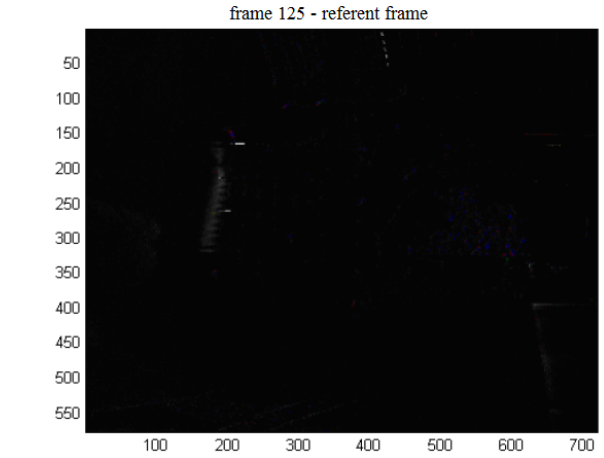
An illustration of the effect of illumination variations on the motion mask is shown in Fig. 5. The video sequence is a recorded operation of an online algorithm without wavelet use, for illustration purposes only. Fig. 5.d shows the entire frame as motion, which is not the correct result.



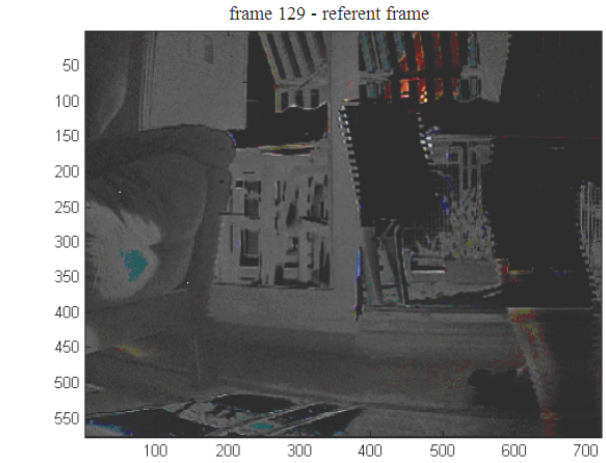
(a)



(b)



(c)



(d)

Figure 5. Illustration of the influence of illumination variations on motion segmentation: (a) frame 125 of the experimental video sequence, (b) frame 129, (c) motion mask for frame 125, (d) motion mask for frame 129

V. CONCLUSION

We investigated a novel approach to motion detection based on energy exponentials. Based on presented research, energy exponentials can be concluded not to produce better results than pure energy algorithms. However, the development of complex robust methods in which the exponential part is used to reduce a part of unwanted effects and energy algorithm to detect total motion, which should be reduced by the energy exponentials' part, has potential. Higher exponentials reduce noise, but are more sensitive to shadows, the waving tree effect and other similar abnormalities, which are unexpected and should be researched in more detail, together with the possible application of the energy exponent algorithms. The results reveal which statistical measures yield better and which give worse results. The results suggest that an algorithm without buffer is less robust to illumination changes (supported by [25]). Another contribution of this research is the development of WEEBD (21), and WCEED (22).

A detailed analysis of results leads us to conclude that different statistical measures do not have the same threshold for the best results. This makes further optimization process an interesting and demanding challenge for further research. Finally, another set of measures could possibly lead us to different conclusions.

REFERENCES

[1] Q. Xie, Q. Long, S. Mita, C. Guo, A. Jiang, "Image Fusion Based on Multi-objective Optimization," International Journal of Wavelets, Multiresolution and Information, vol. 12, no. 2, pp. 1450017, 2014. doi: 10.1142/S0219691314500179

[2] M. Seiferta, H. S. Hock, "The Independent Detection of Motion Energy and Counterchange: Flexibility in Motion Detection," Vision Research, vol. 98, pp. 61–71, 2014. doi: 10.1016/j.visres.2014.03.006

[3] G. T. Zhai, X. L. Wu, X. K. Yang, W. S. Lin, W. J. Zhang, "A Psychovisual Quality Metric in Free - Energy Principle," IEEE Transactions on Image Processing, vol. 21, no. 1, pp. 41-52, 2012. doi: 10.1109/TIP.2011.2161092

[4] D. Y. Huang, T. W. Lin, W. C. Hu, C. H. Cheng, "Gait Recognition Based on Gabor Wavelets and Modified Gait Energy Image for Human Identification," Journal of Electronic Imaging, vol. 22, no. 4, pp. 043039, 2013. doi: 10.1117/1.JEI.22.4.043039

[5] D. J. Joo, "Damage detection and system identification using a wavelet energy based approach," Columbia University, PhD thesis, 2012.

[6] Q. He, "Vibration Signal Classification by Wavelet Packet Energy Flow Manifold Learning," Journal of Sound and Vibration, vol. 332, no. 7, pp. 1881-1894, 2012. doi: 10.1016/j.jsv.2012.11.006

[7] Y. Yang, S. Huang, J. Gao, Z. Qian, "Multi-focus Image Fusion Using an Effective Discrete Wavelet Transform Based Algorithm," Measurement Science and Review, vol. 14, no. 2, pp. 102-108, 2014. doi: 10.2478/msr-2014-0014

[8] K. Y. E. Wong, G. Sainarayanan, A. Chekima, "Palmprint Identification Using Wavelet Energy," in Proc. International Conference on Intelligent and Advanced Systems, Kuala Lumpur, Malaysia, 2007, pp. 714-719, doi: 10.1109/ICIAS.2007.4658480

[9] J. A. Dobrosotskaya, A. L. Bertozzi, "Analysis of the Wavelet Ginzburg-Landau Energy in Image Applications with Edges," SIAM Journal on Imaging Sciences, vol. 6, no. 1, pp. 698–729, 2013. doi:10.1137/100812859

[10] M. Vošvrda, J. Schürrer, "Wavelet Coefficients Energy Redistribution and Heisenberg Principle of Uncertainty," in Proc. Mathematical Methods in Economics, Cheb, Czech Republic, 2015, pp. 894-899.

[11] P. Sebastian, A. Pradeep, "A comparative Study of Artificial Neural Network Based Power Quality Signal Classification Systems with Wavelet Coefficients and Wavelet Based Energy Distribution," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 5, no. 4, pp. 2929-2934, 2016. doi:10.15662/IJAREEIE.2016.0504145

[12] K. Qian, C. Janott, Z. Zhang, C. Heiser, B. Schuller, "Wavelet Features for Classification of Vote Snore Sounds," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, Shanghai, China, 2016, pp. 221-225. doi: 10.1109/ICASSP.2016.7471669

[13] P. K. Bhatia, A. Sharma, "Epilepsy Seizure Detection Using Wavelet Support Vector Machine Classifier," International Journal of Bio-Science and Bio-Technology, vol. 8, no. 2, pp. 11-22, 2016. doi: 10.14257/ijbsbt.2016.8.2.02

[14] S. Y. Elhabian, K. M. E. Sayed, S. H. Ahmed, "Moving Object Detection in Spatial Domain Using Background Removal Techniques - State-of-Art," Recent Patents on Computer Science, vol. 1, no. 1, pp. 32-54, 2008. doi:10.2174/1874479610801010032

[15] S. Manchanda, S. Sharma, "Analysis of computer vision based techniques for motion detection," in Proc. 6th Int. Conf. on Cloud System and Big Data Engineering, Uttar Pradesh, Noida, India, 2016, pp. 445-450. doi:10.1109/CONFLUENCE.2016.7508161

[16] S. Kumar, J. S. Yadav, "Segmentation of moving objects using background subtraction method in complex environments," Radioengineering, vol. 25, pp. 399-408, Jun. 2016. doi:10.13164/re.2016.0399

[17] X. Hu, J. Zheng, "An improved moving object detection algorithm based on Gaussian mixture models," Open Journal of Applied Sciences, vol. 6, pp. 449-456, Jul. 2016. doi: 10.4236/ojapps.2016.67045

[18] M. Shakeri, H. Zhang, "COROLA: A sequential solution to moving object detection using low-rank approximation," Computer Vision and Image Understanding, vol. 146, pp. 27-39, May 2016. doi: 10.1016/j.cviu.2016.02.009

[19] F. Trèves, "Topological Vector Spaces, Distributions and Kernels," pp. 95-126, Academic Press, 1995.

[20] B. Hassibi, A. H. Sayed, T. Kailath, "Indefinite-Quadratic Estimation and Control: A Unified Approach to H2 and H∞ Theories," pp. 81-107, Society for Industrial and Applied Mathematics, 1999. doi: 10.1137/1.9781611970760

[21] N. A. Bruisma, M. A. Steinbuch, "Fast Algorithm to Compute the H∞-Norm of a Transfer Function Matrix," System & Control Letters, vol. 14, no. 4, pp. 287-293, 1990. doi: 10.1016/0167-6911(90)90049-Z

[22] R. Shiavi, "Introduction to Applied Statistical Signal Analysis: Guide to Biomedical and Electrical Engineering Applications," pp. 314, Academic Press, 2007.

[23] Ç. Kocaman, M. Özdemir, "Comparison of Statistical Methods and Wavelet Energy Coefficients for Determining Two Common PQ Disturbances: Sag and Swell," in Proc. International Conference on Electrical and Electronics Engineering, Bursa, Turkey, Nov. 2009, pp. I-80 – I-84 doi: 10.1109/ELECO.2009.5355235

[24] D. Roşca, "Wavelets on Two-dimensional Manifolds," pp. 8, Habilitation thesis, Technical University of Cluj-Napoca, 2012.

[25] I. Vujović, J. Šoda, I. Kuzmanić, "Stabilising Illumination Variations in Motion Detection for Surveillance Applications," IET Image Processing, vol. 7, no. 7, pp. 671-678, 2013. doi: 10.1049/iet-ipr.2013.0169

[26] P. Rosin, E. Ioannidis, "Evaluation of Global Image Thresholding for Change Detection," Pattern Recognition Letters, vol. 24, no. 14, pp. 2345-2356, 2003. doi: 10.1016/S0167-8655(03)00060-6

[27] D. K. Panda, S. Meher, "Detection of Moving Objects Using Fuzzy Color Difference Histogram Based Background Subtraction," IEEE Signal Processing Letters, vol. 23, no. 1, pp. 45-49, 2016. doi: 10.1109/LSP.2015.2498839

[28] S. Davarpanah, F. Khaid, L. Abdullah, "BGLBP-based Image Background Extraction Method," The International Arab Journal of Information Technology, vol. 13, no. 6A, pp. 908-914, 2016.

[29] H. Zhou, G. Su, X. Jiang, "Dynamic Foreground Detection Based on Improved Codebook Model," The Imaging Science Journal, vol. 64, no. 2, pp. 107-117, 2016. doi: 10.1080/13682199.2016.1145313

[30] W. Wang , N. Yang, Y. Zhang, F. Wang, T. Cao, P. Eklund, "A Review of Road Extraction from Remote Sensing Images," Journal of Traffic and Transportation Engineering, vol. 3, no. 3, pp. 271-282, 2016. doi: 10.1016/j.jtte.2016.05.005