Emotion Estimation in Crowds: The Interplay of Motivations and Expectations in Individual Emotions

1st Oscar J. Urizar*,†
*Dept. of Electrical, Electronic, Telecommunications Engineering and Naval Architecture
University of Genoa
Genoa, Italy
oscar.juarez@ginevra.dibe.unige.it

†Department of Industrial Design
Eindhoven University of Technology
Eindhoven, The Netherlands
O.R.Juarez.Urizar@tue.nl

2nd Lucio Marcenaro
Dept. of Electrical, Electronic, Telecommunications Engineering and Naval Architecture
University of Genoa
Genoa, Italy
Lucio.Marcenaro@unige.it

3rd Carlo S. Regazzoni
Dept. of Electrical, Electronic, Telecommunications Engineering and Naval Architecture
University of Genoa
Genoa, Italy
Carlo.Regazzoni@unige.it

4th Emilia I. Barakova
Department of Industrial Design
Eindhoven University of Technology
Eindhoven, The Netherlands
E.I.Barakova@tue.nl

5th Matthias Rauterberg
Department of Industrial Design
Eindhoven University of Technology
Eindhoven, The Netherlands
G.W.M.Rauterberg@tue.nl

***Abstract*—Providing an estimation of the emotional states of individuals increases the insights on the state of a crowd beyond simple normal/abnormal situations or behaviour classification. Methods intended for identifying emotions in individuals are mainly based on facial and body expressions, or even physiological measurements which are not suited for crowded environments as the available information in crowds is usually limited to that provided by surveillance cameras where the face and body of pedestrians can often suffer from occlusion. This work proposes an approach for analysing walking behaviour and exploiting the interplay of motivations and expectations in the emotions of pedestrians. Real-world data is used to test the prediction of motivations and annotations on the emotional state of pedestrians are added to evaluate the proposed method's capability to estimate emotional states. The conducted experiments show significant improvements over previous methods for estimating motivations and consistent results to the estimation of emotions.**

***Index Terms*—pedestrian emotions, emotion estimation, affective models, crowd emotions**

I. Introduction

Field theory describes a person's life to be compound by different fields, and for each field there is a motivation, an attracting force guiding movement towards the goal as well as (potential) repulsive forces preventing to reach the goal [1]. In this sense, motivation is the reason for a person to engage in a given behaviour and interact with an environment (field) to achieve an objective. Consequently, motivations involve expectations [2], a person's assessment of the effort involved to meet a motivation. Emotions play an essential role as they regulate a person's behaviour in an attempt to make reality match expectations [3]. In the context of a crowded environment, a pedestrian engages in a walking behaviour coherent to the intended motivation. The expectation is conceptualised as the desired conditions to fulfil a motivation and the emotion in the individual arises from the departure between expectation and reality, yielding an emotion in the positive spectrum when reality approaches (or surpasses) expectation or an emotion in the negative spectrum for the opposite case.

Incorporating knowledge from the disciplines of force field, probability theory, expectation psychology and emotional theories, the contribution of this work is to propose a data-driven approach capable of inferring the emotion of pedestrians in crowded environments by estimating individual motivations and expectations based on observed walking behaviours, hence, this method is not suited for stationary crowds. The used definition of emotion follows a Dimensional model approach [4] in which emotions are seen as a point in a valence-arousal affect space. However, only the valence component

This work is funded by EACEA Agency of the European Commission under the EMJD ICE PhD program.

is taking into account here. The method employs a dynamic Bayesian network (illustrated in Figure 1) to describe the influence of motivations (as associated to points of interest) and walking behaviour of pedestrians, followed by conceptualising expectations, to finally use these elements to produce an estimation of the emotional state of individual pedestrians.

The organisation of this paper is as follow. A literature review is presented in Section II. Section III introduces the proposed model for predicting motivations, learn expectations and estimate emotional states. In Section IV the capability of the methods is demonstrated using a publicly available dataset. Finally, Section V summarizes the contribution of this work.

II. Literature Review

Previous methods intended for emotion estimation in crowds mainly focus on critical situations such as panic or evacuation scenarios. The authors of [5] and [6] propose ASCRIBE, an agent-based model to describe the mental states (*emotions*, *beliefs* and *intentions*) of individuals in the decision-making process under stressful situations. At the external level it incorporates mechanisms for mirroring mental states between individuals, and at the internal level, it describes how emotions and beliefs affect each other and how both affect a person's intentions. The model was tested by simulations and an empirical study case that compared four models and showed the ASCRIBE model to yield higher prediction accuracy.

The work of [7] proposed multiple methods for estimating emotions based on an evacuation scenario. At its core, these methods learn the topology of the environment to divide it into subregions and make use of a Dynamic Bayesian Network (DBN) to model conditional interactions occurred in each subregion. these interactions are then converted into super states using a Self-organizing map (SOM) and the occurrence of these super states (events) are encoded by a Gaussian mixture model as positive or negative emotions.

Authors in [8] propose a hierarchical Bayesian model in which pedestrians trajectories are used to create a topological map by means of a self-organizing map (SOM), dividing the environment into zones. Pedestrians trajectories are described as a Markov process transitioning between zones, and behaviours are modelled according to the origin and destination; each behaviour is assigned an emotional label (positive, neutral or negative) according to the time required to reach the estimated destination.

As pointed out in [9] a fair comparison between methods is challenging as each one assumes different ways to describe emotions and makes use of simulations or data not publicly available.

III. Methodology

Several theories on emotion agree that emotions arise from the appraisal of a series of internal and external stimulus [3]. However, in a real-world scenario where only the walking behaviour of pedestrians is observable using surveillance cameras, the task of estimating all potential stimulus is challenging. Under the aforementioned settings, this paper focuses on the motivation-related environmental stimuli that have an impact on walking behaviour, proposing a method to infer the emotional state of a pedestrian based on its estimated motivation and expectation. The proposed method starts by defining points of interest (POIs), provided from manual annotations. Based on the existing POIs, direction fields are generated and later incorporated in a bank of Kalman filters in order to produce behaviour models. Motivations correspond to a pedestrian desired to reach a particular POI, for which transition and emission probabilities are defined. Additionally, during the training phase, the mean distance-to-motivation (dtm) for each pair of origin-destination is learned from data and is later used to formulate expectations. During the testing phase, the partially observed trajectory of a pedestrian is compared against the behaviour models to estimate the motivation, followed by computing its expectation. Finally, the actual dtm is compared against the expected dtm, and the measured innovation is used to compute the emotional state of the pedestrian.

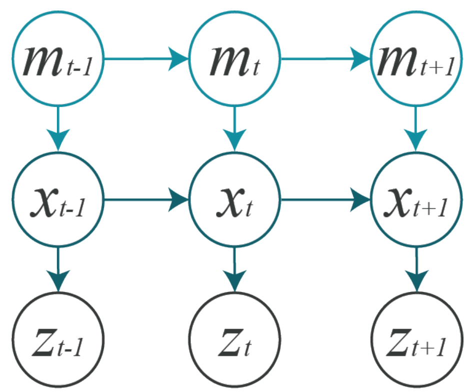


Fig. 1: Dynamic Bayesian network to model the influence of motivation (m_t) on walking behavior (x_t) estimated from observation (z_t).

A. Points of Interest

In the context of this paper, the observed environment contains a discrete set of points of interest (POI) associated with pedestrians' motivations and denoted by the set $\Omega = \{\omega_1, ...\omega_N\}$, where each POI ω_j is associated with an entity (e.g. an object, person, place). In this sense, a POI can be regarded by a particular pedestrian as a point of origin or destination, noting that a POI acts as a motivation (attractor) when considered as a destination. Also, depending on the entity associated, a POI can have a static or dynamic position. A POI is defined as $\omega_j = \{S_j, \vec{D}_j\}$ where $S_j = \{p_1, ..., p_k\}$ is a set of points in $\mathbb{R}^2$ delimiting its bounding box and $\vec{D}_j(p)$ is a radial direction field defined as a vector point-function with the property that for every point in the space of the environment it returns a unit vector pointing towards the area depicted by S_j. Using the points S_j and manual annotations of the environment walls, a direction field $\vec{D}_j$ is generated for each known POI.

ISBN 978-90-827970-1-5 © EURASIP 2018

B. Pedestrian Behavior

The state of pedestrian i at time t is described by its position, indicated with the continuous-valued vector $x_t^i = [x, y]^T$ in $\in \mathbb{R}^2$. Its behaviour is modelled as a Kalman filter, assumed to be influenced by a motivation $m_t^i = j$ where j is the index of the destination $\omega_j \in \Omega$. Hence m^i is treated as a switching variable, with a different behaviour model for each destination

$$x_t^i = F_t x_{t-1}^i + \hat{v}_0^i \vec{D}_j(x_{t-1}^i) + u_k \tag{1}$$

where F_t is the state transition matrix defined as a unit matrix, $\hat{v}_0^i$ is a scalar indicating the estimated desired walking speed, $\vec{D}_j$ the direction field of ω_j, and u_t the interaction forces exerted on pedestrian i at time t modeled as Gaussian white noise with covariance $Q_t : u_t \sim \mathcal{N}(0, Q_t)$. The observation model is given by

$$z_t^i = H_t x_t^i + v_t \tag{2}$$

where H_t is the observation model defined as a unit matrix, and v_t corresponds to the observation noise assumed to be zero-mean Gaussian white noise with covariance $R_t : v_t \sim \mathcal{N}(0, R_t)$. Based on the study presented in [2], this work assumes that a pedestrian tends to maintain the desired walking speed $\hat{v}_0^i$, temporarily affected by its interaction with the environment and other pedestrians. Hence, the desired walking speed is taken from the mean walking speed from previous observations. The innovation with respect to the behaviour model j is defined as

$$\tilde{y}_t^{i,j} = z_t^i - H_t \hat{x}_{t|t-1}^{i,j} \tag{3}$$

where $\hat{x}_{t|t-1}^{i,j}$ corresponds to the estimation produced using $\vec{D}_j$.

C. Pedestrian Motivation

The motivation $m_t^i = j$ describes the intention of pedestrian i at time t to reach the POI ω_j and is conceptualised as a discrete latent variable inferred from its walking behaviour. Therefore is modelled as a Hidden Markov Model (HMM) with N number of states corresponding to the total number of attractors in Ω and transition probability

$$P(m_t^i, \Delta t | m_{t-1}^i) = P(m_t^i | \Delta t, m_{t-1}^i) P(\Delta t | m_{t-1}^i) P(m_{t-1}^i) \tag{4}$$

where the transition probability is conditioned on the previous motivation m_{t-1}^j and the elapsed time Δt until the next transition. The previous motivation m_{t-1}^i is interpreted as the point of origin and m_t^i as the intended destination when $m_{t-1}^i \neq m_t^i$. This work hypotheses ω_j to function as an attractor for pedestrian i when $m_t^i = j$, hence the probability of $m_t^i = j$ given a partially observed path $x_{a:b}^i$ (from observation $t = a$ to $t = b$) is formulated as

$$P(m_t^i = j | x_{a:b}^i) = 1 - \frac{f^j(x_{a:b}^i)}{\sum_{k=1}^{N} f^k(x_{a:b}^i)} \tag{5}$$

where $f^k(x_{a:b}^i) = \sum_{t=a}^{b} \tilde{y}_t^{i,k}$ sums the innovation with respect to the behavior model k as stated in Equation 3 and N is the total number of POIs.

D. Pedestrian Expectation

Research in the field of Expectation Psychology [2] points to the idea that a person engages in a particular behaviour aiming to fulfil a goal by putting an amount of effort close to its psychological expectation. Expectations are formulated according to a person's previous experience, motivation and the context of the environment [10]. This paper uses data from a train terminal. Hence the purpose of pedestrians is predominantly to travel. In this sense, a pedestrian i starts at a location ω_j with the motivation to reach ω_k (i.e. $m_t^i = k$), for which from a first-person perspective there is a distance to travel measured as

$$\mathcal{D}_{actual}^{i,j \to k} = \sqrt{(x_f^i - x_0^i)^2} \tag{6}$$

That is, the Euclidean distance between a pedestrian's initial position x_0^i and $x_f^i \in S_k$ a point of ω_k. The actual time to meet this motivation can be derived as

$$\mathcal{T}_{actual}^{i,j \to k} = \frac{\mathcal{D}_{actual}^{i,j \to k}}{\bar{v}^i} \tag{7}$$

Where $\bar{v}^i$ is the mean walking speed throughout the journey. However, from a pedestrian's perspective, the expected distance is hypothesised (based on observation or previous knowledge) as

$$\mathcal{D}_{expected}^{i,j \to k} = \mathcal{D}_{actual}^{i,j \to k} + c^i \tag{8}$$

Accounting for a perception error, c^i assumed to be zero-mean Gaussian white noise. Consequently, the expected time to reach ω_k is derived from

$$\mathcal{T}_{expected}^{i,j \to k} = \frac{\mathcal{D}_{expected}^{i,j \to k}}{\hat{v}_0^i} \tag{9}$$

with a desired walking speed $\hat{v}_0^i$ chosen according to the level of urgency to meet that motivation and the amount of effort willing to use.

To depict how a pedestrian i with a path x^i approaches towards its desired motivation over time, a motivation-to-distance $dtm^{i,j \to k}$ time series is calculated as described in Algorithm 1. To learn the expected way in which pedestrians with the same origin ω_j and destination ω_k reach their destination, the mean distance-to-motivation $\overline{dtm}^{j \to k}$ of all paths in the group (j, k) is computed. However, as pedestrians walk at different speed hence producing $dtm^{i,j \to k}$ with different arrival times $\tau_{actual}^{i,j \to k}$, the $dtm^{i,j \to k}$ of all pedestrians in the group (j, k) are first adjusted to $\widetilde{dtm}^{i,j \to k}$ using Algorithm 2 where $\overline{\tau}^{j \to k}$ is taken from the mean arrival time. It follows that $\overline{dtm}^{j \to k}$ shows the normality of how a pedestrian approaches its motivation and it will be used in the next section for estimating the emotional state of a pedestrian.

ISBN 978-90-827970-1-5

Algorithm 1 Compute DTM of one trajectory

Input:
$[x^i]$ Pedestrian trajectory
$[x^i_f]$ Final position
Output:
$[dtm^i]$ DTM time series of pedestrian
procedure COMPUTE DTM
  for $k = 1$ **to** $length(x^i)$ **do**
    $dtm^i_k \leftarrow \sqrt{(x^i_f - x^i_k)^2}$

Algorithm 2 Adjust DTM of one trajectory to a desired arrival time

Input:
$[dtm^i]$ DTM time series of pedestrian
$[\tau^{i,j\rightarrow k}_{actual}]$ Actual arrival time of DTM
$[\overline{\tau}^{j\rightarrow k}]$ Adjusted arrival time of DTM
$[\delta_t]$ Size of time interval
Output:
$[\widetilde{dtm}^i]$ Adjusted DTM time series of pedestrian
procedure ADJUST DTM
  $n \leftarrow \tau^{i,j\rightarrow k}_{actual}/\delta_t$, $m \leftarrow \overline{\tau}^{j\rightarrow k}\ \delta_t$
  $dtm\_idx \leftarrow round([1:n] * (m/n))$
  $dtm\_idx_1 \leftarrow 1, dtm\_idx_n \leftarrow m$
  for $t = 1$ **to** m **do**
    $h \leftarrow$ first index value where $dtm\_idx_h >= t$
    if $(t == 1)$ **or** $(t == dtm\_idx_h)$ **then**
      $\widetilde{dtm}_t \leftarrow dtm_h$
    else
      $p = (t - dtm\_idx_{h-1})/(dtm\_idx_h - dtm\_idx_{h-1})$
      $\widetilde{dtm}_t = dtm_{h-1} - (dtm_{h-1} - dtm_h) * p$
    end if
  end for
end procedure

E. Pedestrian Emotion

In the presented model, the emotion of a pedestrian is represented in a single-axis valence with a continuous value ranging from 0 (negative) to 1 (positive). The central idea to estimate the emotional state of a pedestrian is by measuring the deviation between expectation and reality for a particular motivation. Expectation is represented by $\widehat{dtm}^{i,j\rightarrow k}$ which is computed by taking $\overline{dtm}^{j\rightarrow k}$ and adjusting it to an expected arrival time $\tau^{i,j\rightarrow k}_{expected}$ obtained from Equation 9. Reality is given by the actual $dtm^{i,j\rightarrow k}$ computed as time passes. The proportional deviation of reality with respect to expectation is computed by

$$A^i_{deviation} = \frac{A^i_{expected} - A^i_{actual}}{A^i_{expected}} \tag{10}$$

where $A^i_{expected}$ and A^i_{actual} are the trapezoidal area under the curve of $\widehat{dtm}^{i,j\rightarrow k}_{a:b}$ and $dtm^{i,j\rightarrow k}_{a:b}$, respectively, for partial observations from time instance $t = a$ to $t = b$. Also, the value of $A^i_{deviation}$ is constrained to the range $[-1:1]$. Finally, the emotional state $\mathcal{E}^i$ of pedestrian i is computed as

$$\mathcal{E}^i = \epsilon_{expected} + (\epsilon_{expected} * A^i_{deviation}) \tag{11}$$

where $\epsilon_{expected}$ is the expected emotion manually assigned.

IV. EXPERIMENTS AND RESULTS

Experiments are conducted using the Central Station dataset published by [11]. This dataset contains trajectories manually annotated from one-hour-long footage of pedestrians at the main concourse of Grand Central Terminal in New York City. As this dataset does not contain any annotations of the emotional state of pedestrians, nor does any other crowd dataset to the best of our knowledge, the first subsection describes the criteria for generating emotional annotations. The next two subsections evaluate the proposed method's capability to predict motivations and estimate emotional states.

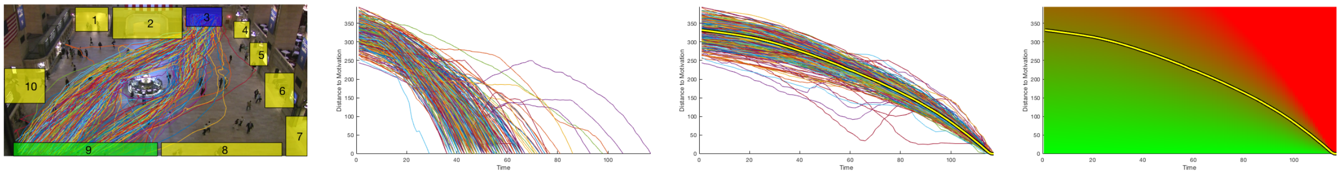
A. Emotional State Annotations

Based on the findings of [2], people in a train station have the motivation of travelling. Motivations can be further divided into commuting, tourism, business, work and others, with each having a different desired waiting time expressed by his or her walking speed. Considering this hypothesis, emotional annotations are obtained by the following process:

(a) Paths are labelled according to their corresponding origin ω_j and destination ω_k and grouped to the pair (j,k).
(b) The $dtm^{i,j\rightarrow k}$ for each path is computed using Algorithm 1 and adjusted to $\widetilde{dtm}^{i,j\rightarrow k}$ with Algorithm 2 using the mean arrival time $\overline{\tau}^{j\rightarrow k}$.
(c) Paths in the group (j,k) are used to compute the mean distance-to-motivation $\overline{dtm}^{j\rightarrow k}$.
(d) For each path, the expected distance-to-motivation $\widehat{dtm}^{i,j\rightarrow k}$ is computed using the estimated desired walking speed $\hat{v}^i_0$ updated at every time instance.
(e) At every time instance t of path i, $A^i_{deviation}$ and $\mathcal{E}^i_t$ are computed by Equations 10 and 11 with $\epsilon_{expected} = 0.5$.

B. Pedestrian Motivation Prediction

The Central station dataset is used to evaluate the proposed method capability to predict a pedestrian's motivation. Similarly to [11], all 12,684 trajectories are included with the first half of each observed trajectory taken as input in the experiment and the methods [11] and [12] are used for comparison. The proposed method is evaluated using different values of θ arbitrarily selected, where θ represents the number of observations used to estimate the motivation. In the case of $\theta = 0$ the innovation $\tilde{y}^{i,k}$ is accumulated for the whole observation period whereas, for $\theta > 0$, the innovation is accumulated using only the last θ number of observations. The experiment's results evidence that including all observations affects performance as pedestrians do not necessarily chose the most optimal path to reach their destination or may change of motivation. Small values of θ tampered accuracy



(a) (b) (c) (d)

Fig. 2: (a) Example of a cluster of tracks with the same origin (coloured blue) and destination (coloured green). (b) Time series of distance-to-motivation for tracks shown in Fig. 2.a. with actual walking speed (c) Time series of distance-to-motivation for tracks shown in Fig. 2.a. normalized to the same walking speed and expected distance-to-motivation in yellow colour. (d) Heat map of the emotional state with valence positive (green) to negative (red) and expected emotion (yellow).

if several POIs are in a similar direction. An optimal value for θ allows for enough evidence to improve accuracy while ignoring potentially misleading segments of a pedestrian path. The results obtained with the proposed method significantly outperform the compared methods [11] and [12].

Method	Parameter	Accuracy
Proposed	$\theta = 0$	66%
Proposed	$\theta = 5$	**73%**
Proposed	$\theta = 10$	71%
Proposed	$\theta = 20$	70%
Proposed	$\theta = 30$	68%
Stationary Crowd [11]	-	48%
MDA [12]	-	43%

TABLE I: Accuracy of motivation prediction. Motivation prediction is computed with Equation 5 using all samples when $\theta = 0$ or using the last θ number of observations.

C. Emotional State Estimation

The estimation of emotion under the proposed hypothesis is subjected to how well the motivation m_t^i and desired walking speed $\hat{v}_0^i$ are estimated. The mean square error (MSE) measurement is adopted to evaluate the emotional state estimation. The results of emotion estimation are presented in Table II, under different values of parameter θ for motivation prediction. Results show a consistent correlation between emotion estimation and motivation prediction as the best results of Table II corresponds to those of $\theta = 5$ which also correspond to the best results in Table I.

θ	0	5	10	20	30
MSE	0.0548	**0.0551**	0.0536	0.0535	0.0538

TABLE II: Mean square error for estimation of emotional state based on motivation prediction using different values of parameter θ.

V. Discussion

This paper introduced a method for estimating individual emotions based on observed walking behaviour and inferred motivations and expectations. The prediction of pedestrian's motivation was addressed using direction fields generated for each POI. The emotional state is derived from the difference between expected and actual observed behaviours. A hypothesis of expectation for pedestrians in a train station was proposed and employed to generate emotional state annotations. The results of the proposed method indicate a significant improvement for predicting motivations (destinations) over previous works, and to efficiently estimate individual emotions based on the proposed hypothesis for expectations. The assumptions made on pedestrians' expectations are crucial for the effectiveness of the method, at the same time, it highlights the adaptability of the method to different environments and types of crowds.

References

[1] B. Burnes and B. Cooke. Kurt Lewin's field theory: A review and re-evaluation. *International Journal of Management Reviews*, 2013.

[2] B. Li, Q.-s. Jin, and H.-y. Guo. Research on Psychological Characteristics of Passengers in Terminal and Some Related Measures. *Procedia - Social and Behavioral Sciences*, 2013.

[3] K. R. Scherer. On the rationality of emotions: or, When are emotions rational? *Social Science Information*, 2011.

[4] J. A. Russell. Core affect and the psychological construction of emotion. *Psychological review*, 2003.

[5] M. Hoogendoorn, J. Treur, C. N. Van Der Wal, et al. Modelling the interplay of emotions, beliefs and intentions within collective decision making based on insights from social neuroscience. *Lecture Notes in Computer Science*, 2010.

[6] T. Bosse, M. Hoogendoorn, M. Klein, et al. Agent-based modelling of social emotional decision making in emergency situations. *Understanding Complex Systems*, 2013.

[7] M. W. Baig, M. S. Baig, V. Bastani, et al. Perception of emotions from crowd dynamics. *International Conference on Digital Signal Processing, DSP*, 2015.

[8] O. J. Urizar, M. S. Baig, E. I. Barakova, et al. A Hierarchical Bayesian Model for Crowd Emotions. *Frontiers in computational neuroscience*, 2016.

[9] O. J. Urizar, E. I. Barakova, L. Marcenaro, et al. Emotion Estimation in Crowds: A Survey. *IET Conference Proceedings*.

[10] T. D. Kemper and R. S. Lazarus. Emotion and Adaptation. *Contemporary Sociology*, 1992.

[11] S. Yi, H. Li, and X. Wang. Understanding pedestrian behaviors from stationary crowd groups. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015.

[12] B. Zhou, X. Wang, and X. Tang. Understanding collective crowd behaviors: Learning a Mixture model of Dynamic pedestrian-Agents. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2012.

ISBN 978-90-827970-1-5 © EURASIP 2018