Automatic Recognition System of Infant Cry based on F-Transform

Nemir Ahmed Al-Azzawi
Baghdad University
Biomedical Engineering Department

ABSTRACT

A new approach has been proposed to design of an automatic infant cry recognition system based on the fuzzy transform (F-transform) that classifies two different kinds of cries, which come from physiological status and medical disease. Feature extraction and development of classification algorithms play an important role in the area of automatic analysis of infant cry signals. F-transform is the powerful tool for the approximation of continuous functions on a finite domain and has been used as feature extraction of infant cry signals. Neural networks is developed to classify the infant cry signals into physiological status and medical disease and trained with smoothing parameter to obtain better classification accuracy. Neural networks (recognition system) are trained from the random selection of crying babies. The experimental results show that the proposed give very promising classification accuracy of 96%. The proposed method can be used to assist medical investigation of the infant from the cry signals.

Keywords
Fuzzy-Transform (F-Transform); Infant cry; Artificial neural network; Automatic recognition; Feature extraction.

1. INTRODUCTION
Infant crying is the only way and the mainly influential to express their feelings. It has been confirmed that the acoustic characteristics of the cry sound are straight influenced by the infant's physical and psychological state or various external stimuli [1]. Special characteristics of newborn pain cries have been identified in various pathological states [2]. The pathological diseases in infants are commonly detected several months or years, after the infant is born. If any of these diseases could be detected earlier, they can be attended and perhaps avoided by the opportune application of treatments and therapies [3]. Thus, great advantage of having automatic classification is that infant baby cry indicator will be independent of the domain expert. Additionally, clinician can obtain an early diagnostic result if the baby having high probability of getting various type of medical diseases.

A system of this type has been suggested by Cohen and Zmora [4], and was shown to be successful in classification of a preliminary database that consists of hunger and pain cries of healthy full-term infants. The system used a simple minimum distance classifier. There are many studies that conducted continuously from the baby age 1 month to 1 year old [5-8]. Moller and Schonweiler described the early detection of hearing impairment [9]. John describe the neural structures underlying cry production, or the circuitry that mediates a caregiver's response to cry sounds [10]. Radhika estimated the fundamental and formant frequencies of infants cries with congenital heart disorder using frequency domain (Cepstrum) and linear prediction code (LPC) methods [11]. Cry is a common response that being investigated related to

pain [12-14], hungry [13-15], no-pain, discomfort [1], apnea [2], asphyxia [16-20], hypothyroidism [19, 21], Hyperbilirubinemia [22], cleft palate [13, 23], Ankyloglossia [24], respiratory distress syndrome [13], hearing disorder [25], brain damage [3], hydrocephalus [26] and sudden infant death syndrome (SIDS) [27].

Instead of wavelet transform [28] or linear prediction technique [29], F-transform method [30] has been suggested to extract the features of infant cry. This method is investigate suitability for the problem of infant cries classification. In [30], it has been shown that fuzzy transform is the powerful tool for approximation of continuous functions on a finite domain. The classification of baby cry has been intended by the use neural networks recognition approaches.

In this paper, the design and implementation of a system that classifies two different kinds of cries physiological status and medical disease are discussed based on F-transform feature extracted. Neural network has been trained by using a randomly selected set of infant cry samples and establish an automatic recognition system of infant cry. The experimental results demonstrate the robustness, efficiency, encouraging and accuracy of the proposed algorithm. The results provide additional support for the contention that these infants take longer to produce and demonstrate a noticeable decrease.

2. BASIC CONCEPT OF F-TRANSFORM
The original motivation for the F-transform (an abbreviated name for the fuzzy transform) came from fuzzy modeling Perfilieva [31]. The intention was to show that, similarly to conventional transforms (Fourier and wavelet). Let u be represented by the discrete function $u : P \rightarrow \mathbb{R}$ of one variables, where $P = \{(i) \mid i = 1, \ldots, N\}$ is an N vector of discrete speech signal, and $\mathbb{R}$ is the set of reals. If $i \in P$ is a sample, then $u(i)$ represents its intensity range. Moreover, let fuzzy sets A_k $k=1,\ldots,n$, where $0 < n \leq N$, establish a fuzzy partition of $[1, N]$.

The F-transform of u corresponds u to the matrix $F[u]_n$ of F-transform components:

$$F[u]_n = (F[u]_1 \ldots F[u]_n) \qquad (1)$$

Each component $F[u]_k$ is a local mean value of u over a support set of the respective fuzzy set A_k. The membership functions of the respective fuzzy sets in a fuzzy partition are called basic functions. The (direct) F-transform of u (with respect to the chosen partition) is an speech signal represent $F[u] : \{A_1,\ldots, A_n\} \rightarrow \mathbb{R}$ defined by:

$$F[u](A_k) = \frac{\sum_{i=1}^{N} u(i) A_k(i)}{\sum_{i=1}^{N} A_k(i)} \qquad (2)$$

where k = 1,... , n,. The value $F[u](A_k)$ is called an F-transform component of u and is denoted by $F[u]_k$. The

extreme fuzzy partitions of [1, *N*] are introduce, that will be used in the following.

Largest partition: The largest partition of [1,*N*] contains only one fuzzy set, A_1, such that for all x∈[1,*N*], (A_1)(x)=1. The respective F-transform component $F[u]_1$ and the respective inverse F-transform u_1.

Finest partition: The finest partition of [1, *N*] is established by *N* fuzzy sets A_k, such that for all *k*=1,...,*N*. The respective F-transform components $F[u]_k$, *k*=1,...,*N*, and the respective inverse F-transform u_N. The inverse F-transform of *u* is a function on *P*, which is represented by the following inversion formula:

$$u_n(i) = \sum_{i=1}^{n} F[u]_k A_k(i) \tag{3}$$

where *i* = 1,... , *N*. It can be shown that the inverse F-transform, u$_n$ approximates the original function *u* on the domain *P*. The proof can be found in [30, 32]. The F-transform technique, leading to one-level or higher-level decomposition of a discrete speech signal; here the explain the technical details of these decompositions are shown. The one level decomposition is as the following representation of *u* on *P*:

$$u(x) = u_n(x) + e(x) \tag{4}$$

$$e(x) = u(x) - u_n(x), \forall(x) \in P \tag{5}$$

where 0<*n*≤*N* and u_n is the inverse F-transform of *u* and *e* is the respective residuum. If function is smooth, then the error function $e(x)$ is small, and the one-level decomposition (5) is sufficient for our feature extraction algorithm. Conversely, discrete speech signal generally have various types of degradation that interrupt their smoothness. As a result, the error function $e(x)$ in (5) is not negligible, and the one-level decomposition is insufficient for our purpose. Therefore, continue with the decomposition of the error function *e* in (5).

This will lead to decompose into its inverse F-transform $e_{n'}$ (with respect to a finer fuzzy partition of [1,*N*] with $n':n<n'≤N$ basic functions, respectively) and a new error function e'. Thus, the second-level decomposition of *u* are obtained. Note that the number of fuzzy sets are possibly increase in which to create fixed fuzzy transform arbitrarily and convergence is uniform. Also the starting approximation can be taken as fixed fuzzy transformation of a higher level (number of fuzzy sets > 1). Looking at the complexity of fuzzy transform, it is $O(md2^d)$ (*d* - dimension, *m* - data size) that is the complexity of the same order as in the case of wavelet transform [30].

3. THE PROPOSED F-TRANSFORM APPROACH

The classification algorithm is based on the F-transform. The first step of the automatic processing system is segmentation of the sampled cry signals to cry and silent segments. Whereas, recorded infant cries in the databases used in a reasonably high signal-to-noise ratio, simple thresholds on the signal energy envelope are adequate to perform segmentation. The energy is computed over 10 milliseconds of the signal every 10 milliseconds in order to obtain the energy envelope of the signal. Two thresholds are used as a certain percentage of the average energy over 10 seconds of signal. The first

threshold, which is the highest, is used to search the energy curve for cry segments, meaning, to distinguish among cry and silent segments, whereas the second threshold is used to adjust the endpoints of the cry segments. Just cry segments longer than a minimal duration of 60 milliseconds are processed.

The features extracting is very important as it concerns the complexity of the neural network and the precision of the speech recognition. According to the F-transform approach, the feature extraction uses as follows:

1. Compute $F[u]$ – the direct F-transform of discrete signal $u : P \rightarrow \mathbb{R}$, by (2).

2. Compute u_n – the inverse F-transform using the components $F[u]$ by (3).

3. Compute the error function $e(x) = |u(x) - u_n(x)|$ for all $x \in P$.

4. Rescale and round the values of *e* from [0, $max_{x \in P} e(x)$], which results in the new e_r.

These features are used as an input to the neural network in the next block for classification. The task of the classification system is to automatically distinguish between given types of signals. The different types of cry signals are used, for physiological status, medical disease 1, medical disease 2, etc.

Neural network is a technology which tries to mimic human brain functions. With the improvement of neural network these past few decades, using neural network in speech and speaker recognition has become very popular and successful [33].

The classification system automatically classifies the specified signal to one of the types. Supervised classification systems, such as in our case, consists of two phases: a training stage and a recognition stage. The training is usually a "one time" task, but after the initial training, an adaptation scheme may be used to adjust the models for varying conditions.

The used classifier is multi-layer perceptron (MLP), neural network, for training process, the cry sample by 256 elements in feature vector are more require. On other side with size of this feature vector, computational cost is high. To overcoming to this problem, the principle components analysis (PCA) are used. PCA analysis reduced size of feature vector from 256 to 32 elements, so ANN can be trained in lower cost. The neural networks participate several characteristics, these properties are the basis to configure any neural network, and they are, computational unites (nodes), connections between nodes and training the networks[34].

Architecture of our used MLP consists of 32 input layers, 6 nodes in hidden layer and 2 nodes in output layer and conjugate gradient method [35] was used for network training. The recognition procedure calls for the presentation to the classification system of a sample of the cry.

Figure 1, shows a block diagram of the proposed automatic recognition system of infant cry based on F-transform. Finally the decision is taken in end stage. Where, the models are estimated; their parameters are placed in memory and employed by the recognition system.
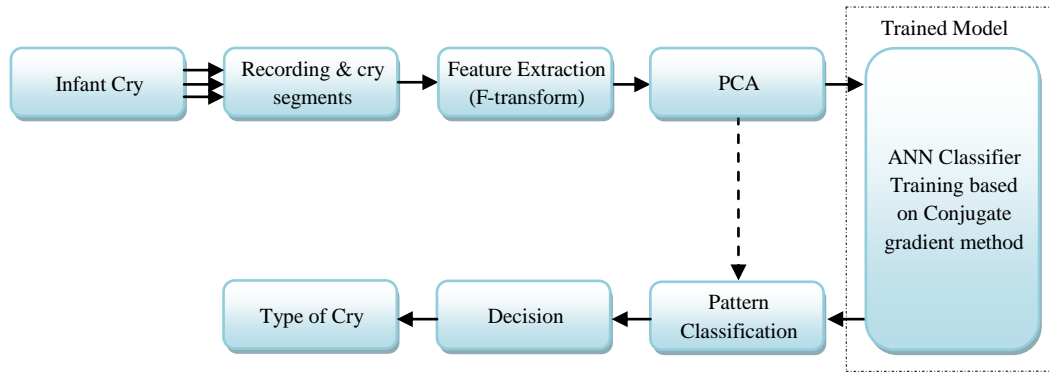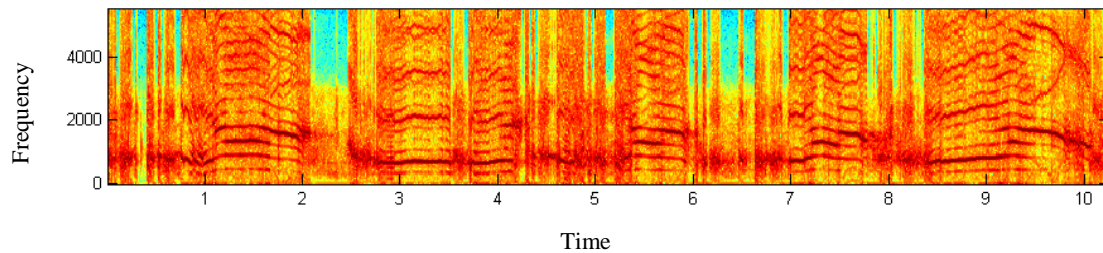
Fig. 1: Block diagram of the proposed automatic recognition of infant cry based on F-transform.

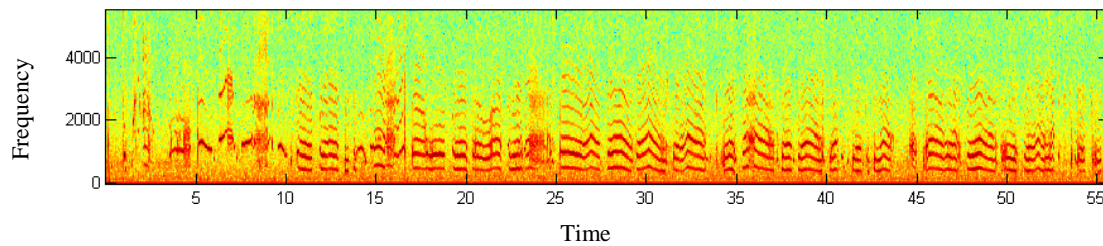(a) Physiological status

(b) Medical disease

Fig. 2: The spectrogram displays of (a) the crying of a physiological status infant and (b) infant with medical disease.

4. EXPERIMENT AND RESULTS ANALYSIS

To test our algorithm, two types of crying were selected: physiological status and medical disease. The database of infant cry babies has been collected from several clinics in my country, and also from certain cry sample database: Instituto Nacional de Astrofisica Optica y Electronica (INAOE) [36]. The cry data set is made up of 276 cry samples records, from both categories, with different duration. The cry data signal were segmented each signal to segments with 1 s of length. 836 segmented samples were obtained, 620 of them belong to physiological status, and 216 to medical disease. All signals were normalized to 16 kHz sampling rate and 16 bit quantized. F-transform is the powerful tool for the approximation of continuous functions on a finite domain and has been used as feature extraction of infant cry signals. The used classifier is 'multi-layer perceptron (MLP)' neural network, for training process. Table 1, shows the accuracy in the classification, which is 96 % based on F-transform.

Table 1. Automatic recognition system of infant cry based on F-transform.

Type of crying	Samples	Disorder Decision	
		Physiological	Medical
Physiological	620	598	22
Medical disease	216	8	208
Total	836	Accuracy is 96%	

According to results in Table I, shows that F-transform features are efficient features in classification of infant cry. The automatic recognition systems of infant cry based on F-transform were implemented in Matlab including preprocessing, F-transform, neural network and PCA. The neural network and the training method were done with the neural networks toolbox of Matlab. The implementation of PCA technique employed is as described by [37].

An experimental result is compared with automatic recognition system of infant cry based on linear prediction technique [29]. Table 2, shows the accuracy in the classification, which is 85 % based on linear prediction technique.

It can be easily seen that accuracy of recognition of proposed method has better results than results compared with previous method based on linear prediction technique [29]. Figure 2, shows the spectrogram displays of the crying of a physiological status infant (part A) and an infant with medical disease (part B).

Table 2. Automatic recognition system of infant cry based on linear prediction technique.

Type of crying	Samples	Disorder Decision	
		Physiological	Medical
Physiological	620	527	93
Medical disease	216	33	183
Total	836	Accuracy is 85%	

5. CONCLUSION

Evaluation of newborn cry is a powerful viewing tool to evaluate, manage and provides an early window into the neurological status. Atypical cries can be viewed as a helpful monitor that should be referred for a full neurological work up. In this paper a new automatic recognition system of infant cry has been proposed based on new feature extraction F-transform. After comparison of accuracy recognition system obtain from (F-transform based) and (linear prediction based), experimental results showed that proposed method is efficient in automatic recognition system of infant cry. The proposed could better recognize infant's cry with extraction of F-transform feature. Therapy with parents to recognize their infants who illustrate atypical cry characteristics can assist a helpful developmental context during in fancy and young childhood. Future directions can to develop a portable, easy to use cry collection device integrated with analysis software that utilizes algorithms for the infant vocal tract and calculates outcome measures. Incorporate measurement of cry perception in high risk mothers or with high risk infants in primary care, early intervention, and occupational therapy. This information will be very helpful to pediatricians and doctors in general.

REFERENCES

[1] R. E. Stark, *et al.*, "Features of infant sounds: The first eight weeks of life," *Journal of Child Language,* 2(02), pp. 205-221, 1975.

[2] M. P. Robb, *et al.*, "Cry analysis in infants resuscitated for apnea of infancy," *International Journal of Pediatric Otorhinolaryngology,* 71(7), pp. 1117-1123, 2007.

[3] B. M. Lester and C. F. Zachariah Boukydis. (1985). *Infant Crying Theoretical and Research Perspectives* [Online]. Boston, MA: Springer US. Available: http://dx.doi.org/10.1007/978-1-4613-2381-5 MIT Access Only [Accessed].

[4] A. Cohen and E. Zmora, "Automatic classification of infants' hunger and pain cry," in *Proc. Int. Conf. Digital Signal Processing, Cappellini, V. and Constantinides, AG, Eds., Elsevier, Amsterdam*, 1984, pp. 667-672.

[5] H. R. Gilbert and M. P. Robb, "Vocal fundamental frequency characteristics of infant hunger cries: birth to 12 months," *International Journal of Pediatric Otorhinolaryngology,* 34(3), pp. 237-243, 1996.

[6] L. Rautava, *et al.*, "Acoustic quality of cry in very-low-birth-weight infants at the age of 1 1/2 years," *Early Human Development,* 83(1), pp. 5-12, 2007.

[7] T. Etz, *et al.*, "Infant cry reliability: Acoustic homogeneity of spontaneous cries and pain-induced cries," *Speech Communication,* 58(0), pp. 91-100, 2014.

[8] A. K. Singh, *et al.*, "Classification of infant cries using epoch and spectral features," in *Communications (NCC), 2013 National Conference on*, 2013, pp. 1-5.

[9] S. Möller and R. Schönweiler, "Analysis of infant cries for the early detection of hearing impairment," *Speech Communication,* 28(3), pp. 175-193, 1999.

[10] D. N. John, "Review Neural circuits underlying crying and cry responding in mammals," 2013.

[11] L. Radhika Rani, *et al.*, "Estimation of Fundamental and Formant frequencies of infants ' cries; a study of Infants with congenital Heart disorder," 2013.

[12] P. Runefors, *et al.*, "Newborn infants' cry after heel-prick: analysis with sound spectrogram," *Acta Paediatrica,* 89(1), pp. 68-72, 2000.

[13] D. Lederman, *et al.*, "On the use of hidden Markov models in infants' cry classification," in *Electrical and Electronics Engineers in Israel, 2002. The 22nd Convention of*, 2002, pp. 350-352.

[14] S. E. Barajas-Montiel and C. A. Reyes-Garcia, "Identifying pain and hunger in infant cry with classifiers ensembles," in *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on*, 2005, pp. 770-775.

[15] H. E. Baeck and M. N. de Souza, "Longitudinal study of the fundamental frequency of hunger cries along the first 6 months of healthy babies," *Journal of Voice,* 21(5), pp. 551-559, 2007.

[16] O. F. R. Galaviz and C. A. R. Garcia, "Infant Cry classification to identify hypoacoustics and asphyxia with neural networks," in *MICAI 2004: Advances in Artificial Intelligence*, ed: Springer, 2004, pp. 69-78.

[17] O. F. R. Galaviz and C. A. R. García, "Infant cry classification to identify hypo acoustics and asphyxia comparing an evolutionary-neural system with a neural network system," in *MICAI 2005: Advances in Artificial Intelligence*, ed: Springer, 2005, pp. 949-958.

[18] A. Zabidi, *et al.*, "Classification of Infant Cries with Asphyxia Using Multilayer Perceptron Neural Network," in *Computer Engineering and Applications (ICCEA), 2010 Second International Conference on*, 2010, pp. 204-208.

[19] A. Zabidi, *et al.*, "Detection of infant hypothyroidism with mel frequency cepstrum analysis and multi-layer perceptron classification," in *Signal Processing and Its Applications (CSPA), 2010 6th International Colloquium on*, 2010, pp. 1-5.

[20] M. Hariharan, *et al.*, "Analysis of infant cry through weighted linear prediction cepstral coefficients and probabilistic neural network," *Journal of Medical Systems,* 36(3), pp. 1309-1315, 2012.

[21] K. Michelsson and P. Sirviö, "Cry analysis in congenital hypothyroidism," *Folia Phoniatrica et Logopaedica,* 28(1), pp. 40-47, 1976.

[22] B. R. Vohr, *et al.*, "Abnormal brain-stem function (brain-stem auditory evoked response) correlates with acoustic cry features in term infants with hyperbilirubinemia," *The Journal of Pediatrics,* 115(2), pp. 303-308, 1989.

[23] D. Lederman*, et al.*, "Classification of cries of infants with cleft-palate using parallel hidden Markov models," *Medical & biological engineering & computing,* 46(10), pp. 965-975, 2008.

[24] Y. Okada*, et al.*, "Iterative forward on cross-validation approach and its application to infant cry classification," in *International Multi Conference of Engineers and Computer Scientists*, Hong Kong, 2011.

[25] G. Várallyay Jr*, et al.*, "Acoustic analysis of the infant cry: classical and new methods," in *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, 2004, pp. 313-316.

[26] K. Michelsson*, et al.*, "Crying in separated and non-separated newborns: sound spectrographic analysis," *Acta Paediatrica,* 85(4), pp. 471-475, 1996.

[27] R. Colton and A. Steinschneider, "The cry characteristics of an infant who died of the sudden infant death syndrome," *Journal of Speech and Hearing Disorders,* 46(4), pp. 359-363, 1981.

[28] Z. Tufekci*, et al.*, "Applied mel-frequency discrete wavelet coefficients and parallel model compensation for noise-robust speech recognition," *Speech Communication,* 48(10), pp. 1294-1307, 2006.

[29] J. Orozco and C. A. R. García, "Detecting pathologies from infant cry applying scaled conjugate gradient neural networks," in *European Symposium on Artificial Neural Networks, Bruges (Belgium)*, 2003, pp. 349-354.

[30] I. Perfilieva, "Fuzzy transforms: Theory and applications," *Fuzzy Sets and Systems,* 157(8), pp. 993-1023, 2006.

[31] I. Perfilieva*, et al.*, "Fuzzy transform in the analysis of data," *International Journal of Approximate Reasoning,* 48(1), pp. 36-46, 2008.

[32] I. Perfilieva, "Fuzzy transforms: A challenge to conventional transforms," in *Advances in Images and Electron Physics.* vol. 147, ed San Diego: Elsevier Academic Press, 2007, pp. 137-196.

[33] R. G. Barr*, et al.*, *Crying as a sign, a symptom, & a signal : clinical, emotional, and developmental aspects of infant and toddler crying*. London: Mac Keith Press : Distributed by Cambridge University Press, 2000.

[34] R. Rojas, *Neural networks : a systematic introduction*, 1996.

[35] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization,* 16(1), pp. 170-192, 2005.

[36] C. A. R. GARCÍA. (2003). *Detecting Pathologies from Infant Cry* [Online]. Puebla, México: Instituto Nacional de Astrofisica Optica y Electronica (INAOE). Available: http://ccc.inaoep.mx/~kargaxxi/Archivos/onesecod/ [Accessed].

[37] N. Al-Azzawi*, et al.*, "Medical image fusion scheme using complex contourlet transform based on PCA," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2009*, Hilton Minneapolis, Minnesota, USA, 2009, pp. 5813-5816.