This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

Medical Image Analysis Using AM-FM Models and Methods

Kyriacos P. Constantinou, Ioannis P. Constantinou, Constantinos S. Pattichis, *Fellow, IEEE,* and Marios S. Pattichis, *Senior Member, IEEE*

Methodological Review

***Abstract*—Medical image analysis methods require the use of effective representations for differentiating between lesions, diseased regions, and normal structure. Amplitude Modulation - Frequency Modulation (AM-FM) models provide effective representations through physically meaningful descriptors of complex non-stationary structures that can differentiate between the different lesions and normal structure.**

Based on AM-FM models, medical images are decomposed into AM-FM components where the instantaneous frequency provides a descriptor of local texture, the instantaneous amplitude captures slowly-varying brightness variations, while the instantaneous phase provides for a powerful descriptor of location, generalizing the traditionally important role of phase in the Fourier Analysis of images.

Over the years, AM-FM representations have been used in a wide variety of medical image analysis applications based on a vastly reduced number of features that can be easily learned by simple classifiers. The paper provides an overview of AM-FM models and methods, followed by applications in medical image analysis. We also provide a summary of emerging trends and future directions.

***Index Terms*—Image representations, AM-FM models, medical image analysis, texture.**

I. Introduction

The complex non-stationary characteristics of medical image structures require the development of effective, easy to visualize, and meaningful image representations. Medical image regions can often be characterized by complex textures that are associated with specific diseases. There is a clear need for medical image models that can be spatially adaptive, physically meaningful, and that can also be easily modified to target specific diseases.

The recent development of Convolutional Neural Networks has provided methods that yield excellent segmentation and classification results for several image analysis methods (e.g., see [1]). The lower layers of Convolutional Neural Networks are defined in terms of the use of several layers of interconnected filterbanks, followed by max pooling layers to extract image features. Yet, due to the large number of layers that are often involved, Convolutional Neural Networks are very hard to interpret, requiring a careful analysis of the outputs of a large numbers of convolutional layers. Unfortunately, there is no clear path for analyzing convolutional layers in the frequency domain because linearity is lost due to the use of several max pooling layers. Furthermore, if not required, interpretability is highly desired for medical imaging applications. The lack of clear interpretability of Convolutional Neural Networks can significantly limit their wider acceptance by medical professionals. For applications in medical diagnosis, explainability is required by law in the European Union [2].

AM-FM models are inherently interpretable by using well understood concepts based on the instantaneous frequency and the instantaneous amplitude. When combined with filterbanks, a primary motivation for the current paper, AM-FM models provide a simple and natural interpretation of the filterbank outputs. AM-FM filterbanks are directly defined in the frequency domain while their outputs are naturally described using AM-FM functions. AM-FM methods use a single filterbank layer defined using a small number of parameters (e.g., number of orientations and a half-peak overlap requirement) that can be easily visualized in the frequency domain. There is clearly strong motivation to study AM-FM representations for understanding filterbank outputs and also to use AM-FM representations for learning interpretable filterbanks that are used with Convolutional Neural Networks, extending ongoing research (e.g., see [3], [4]).

AM-FM methods can clearly benefit from the use of deep learning methods for feature selection and classification. As we shall describe later in this paper, AM-FM representations generate several AM-FM components with associated features (e.g., instantaneous frequency and instantaneous amplitude features) that require feature selection. Deep learning methods can be used to support feature selection and classification using interpretable AM-FM features. Yet, so far, there has not been a major study that combines AM-FM features with deep learning. Deep learning methods can be used to reduce the large number of generated AM-FM features and then use them to provide better classification performance.

We proceed to introduce AM-FM models and explain how they can be used to provide effective representations for biomedical images. Let $I(x,y)$ denote an input image. We decompose the image into its AM-FM components using (e.g., see early work in [5]):

$$I(x,y) = \sum_{n=1}^{K} a_n(x,y) \cos \varphi_n(x,y) \tag{1}$$

K. P. Constantinou, I. P. Constantinou and C. S. Pattichis are with the Department of Computer Science, University of Cyprus, Nicosia, Cyprus. They can be reached by email at: csp7ck1@cs.ucy.ac.cy, ioannis@istognosis.com, and pattichi@cs.ucy.ac.cy.

M. S. Pattichis is with the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM-87131, USA. E-mail: pattichi@unm.edu.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

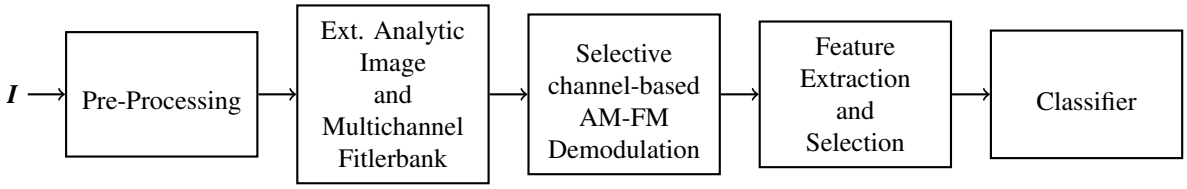

Fig. 1. A framework for computing multiscale AM-FM models and features for medical image classification.

where $n = 1, \ldots, K$ indexes the AM-FM components, a_n denotes the instantaneous amplitude (IA) for the n-th AM component, and φ_n denotes the instantaneous phase (IP) for the n-th FM component. In addition, the gradient of the phase, $\nabla\varphi_n(x, y)$ represents the instantaneous frequency (IF) for the n-th FM component.

For medical image analysis applications, the IA is associated with the spatial extend of specific structures (e.g., see [6], [7]). To see this, first note that AM-FM decompositions are often computed through the application of a filterbank. Hence, since convolution is a linear operator, brighter regions will clearly generate stronger responses and hence large values for the corresponding $a_n(x, y)$. On the other hand, darker regions will then generate weaker responses and lower values for the corresponding $a_n(x, y)$. Furthermore, the IA is slowly-varying as opposed to strong spatial texture variations that generate high-frequency components [6], [8], [7]. Complex texture variations are described by the FM components $\cos\varphi_n(x, y)$ (see [8]).

AM-FM representations allow for an intuitive, spatially adaptive interpretation of local frequency content. Locally, the IF vectors are orthogonal to ridges or edge structures and have a magnitude that is inversely proportional to the instantaneous wavelength, the distance to the next ridge or strong edge [9]. More generally, FM flow lines can be used to describe changes throughout the image as described by the ordinary differential equations derived in [9]. In medical image analysis applications, we often find that discriminating features are associated with specific ranges of values of the IF components. In particular, to maintain rotational invariance in the extracted features, we extract AM-FM components with different ranges of the IF magnitude. In particular, we use the terms: (i) low-scale, (ii) medium-scale, and (iii) high-scale to refer to low, medium, and high magnitudes of the IF.

We introduce a typical framework for computing AM-FM decompositions in Fig. 1. Initially, the image can be preprocessed for normalization, to perform illumination correction, and to standardize image resolution. Image normalization and illumination correction will help reduce non-uniform illumination artifacts and hence reduce variability in the extracted AM-FM features. If possible, the standardization of image resolution refers to the application of image resampling methods to ensure that all images are sampled at the same number of pixels per unit length (e.g., pixels per mm). It is important to note that the standardization of image resolution will ensure that the extracted instantaneous frequencies will correspond to actual physical frequencies. The preprocessed images are then analytically extended and filtered through a multichannel filterbank. For computing multiscale decompositions, we associate scales with specific groups of filters (e.g., low, medium, or high frequency magnitude filterbank channels). The AM-FM components and features are then extracted from each scale or selected group of filters and input into a classifier.

We present a multiscale example in Fig. 2. In this example, the goal is to differentiate between normal and abnormal regions in mammography. By comparing the original input images (see Figs. 2(a) and 2(d)), we can see that the abnormal region appears brighter. Yet, in terms of the IA PDF plots of Figs. 2(g) and 2(i), we note that the opposite pattern is revealed for the medium and high scales. The IA for the abnormal regions for these scales are much lower, suggesting that the original image brightness reflects the behavior at the lowest scales. From the medium-scale AM-FM reconstructions of Figs. 2(b) and 2(e), we can see a nearly binary appearance of the abnormal region as opposed to the grayscale appearance of the normal region. The dramatic difference is due to the fact that the PDF of the IA of the abnormal region is far more concentrated near the peak as opposed to the PDF of the IA that is widely spread (see Fig. 2(g)). In terms of the IF, we can see the presence of higher IF magnitudes for the abnormal region at the high scale (see Figs. 2(j), (c), (f)).

The objective of this paper is to present a comprehensive review of application of multiscale AM-FM models and methods in medical imaging. The review is far more exhaustive, detailed, up to date and significantly different than an earlier review of emerging AM-FM methods presented in [10]. The coverage includes a thorough discussion of AM-FM representations, extensive coverage of demodulation methods, and an expanded list of medical imaging applications, including several applications that appeared since the publication of [10].

In Section II, we provide a summary of continuous-space and discrete-space AM-FM representations. In Section III, we provide a summary of methods for estimating AM-FM decompositions. We then provide a comparison of different methods in Section IV and provide several applications in medical image analysis in Section V. Emerging trends possible future directions are summarized in Section VI. We provide concluding remarks in Section VII.

II. AM-FM Representations

In this section, we describe AM-FM representations that can be derived from popular Fourier representations and filterbanks. We begin with continuous-space representations in section II-A. We summarize discrete-space representations in section II-B. Multiscale AM-FM representations are presented in section II-C.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

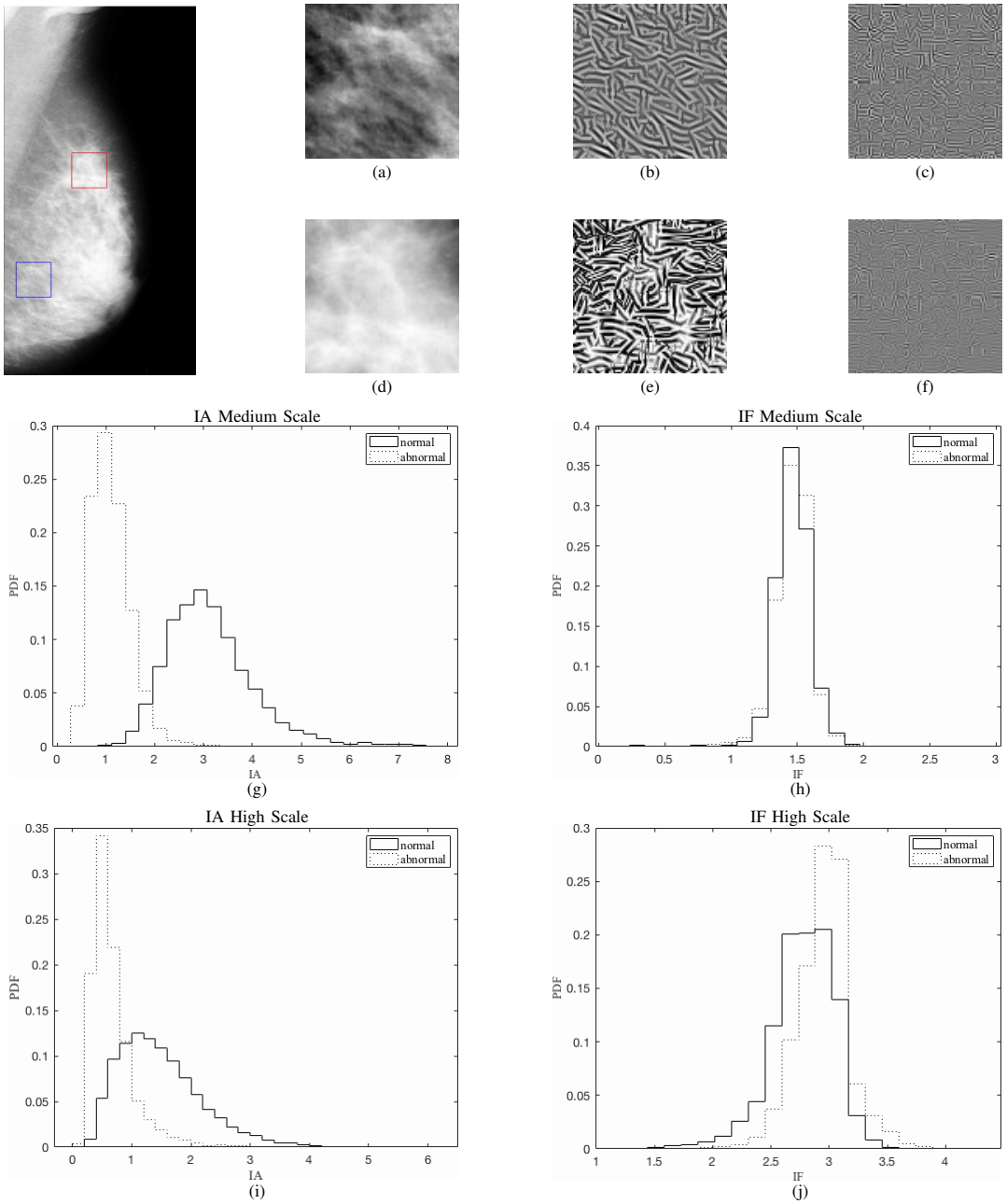

Fig. 2. Mammography image analysis example. The example demonstrates the separation of normal from abnormal ROIs using multiscale AM-FM methods. AM-FM reconstructions are of the form $a(x,y)\cos\varphi(x,y)$. The IF magnitude is given by $||\varphi(x,y)||$. All images have been normalized for maximum contrast. (a) Normal ROI. (b) Medium-scale AM-FM reconstruction of normal region. (c) High-scale AM-FM reconstruction of normal region. (d) Abnormal ROI. (e) Medium-scale AM-FM reconstruction of abnormal region. (f) High-scale AM-FM reconstruction of abnormal region. (g) Medium-scale Probability density function (PDF) plots for normal IA using a continuous line versus abnormal IA using a dotted line. (h) Medium-scale PDF plots for normal IF magnitude using a continuous line and abnormal IF magnitude using a dotted line. (i) Same as in (g) for high-scale. (j) Same as in (h) for high-scale.

A. Continuous-Space AM-FM Representations Derived from Fourier Representations

Efficient continuous-space AM-FM representations can be derived from Fourier-domain representations. As we describe next, the basic idea is to extend the application of effective Fourier representations from an idealized image space to the projected space of observed images.

Let I denote the observed image and I_{ideal} denote an ideal image model that has an effective representation in the Fourier domain. We model the observed image in terms of the ideal image using $I(x,y) = a(x,y) \cdot I_{\text{ideal}}\left(\phi_{co1}(x,y), \phi_{co2}(x,y)\right)$ where the amplitude function $a(x,y)$ captures image illumination variations and $(x,y) \rightarrow \Phi(x,y) = (\phi_{co1}(x,y), \phi_{co2}(x,y))$ describes a curvilinear coordinate transformation associated with the resampling of the ideal image as the observed image.

An early medical imaging example was presented in [6]. In [6], the authors developed an AM-FM model that describes the regular spacing of sarcomeres in human skeletal muscle as observed in electron microscopic images. The basic idea

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

TABLE I
SUMMARY OF NOTATION USED THROUGHOUT THE PAPER. WE USE $x-y$ FOR CONTINUOUS-SPACE COORDINATES (E.G., $I(x,y)$) AND k_1-k_2 FOR DISCRETE-SPACE COORDINATES (E.G., $I[k_1,k_2]$). WE ALSO USE $\hat{I}$ FOR FOURIER TRANSFORMS AND $\check{I}$ FOR AM-FM TRANSFORMS.

Continuous-space AM-FM Representations **Symbol**	**Definition**
$I(x)$	1D signal defined over x coordinate.
$I(x,y)$	image defined on $x-y$ coord. system.
$\hat{I}(f_1,f_2)$	2D Fourier Transform of $I(x,y)$.
$\check{I}(f_1,f_2)$	2D AM-FM Transform of $I(x,y)$.
$a_n(x,y)$	n-th Instantaneous Amplitude comp.
$\varphi_n(x,y)$	n-th Instantaneous Phase component.
$\nabla\varphi_n(x,y)$	n-th Instantaneous Frequency comp.
$\cos\varphi_n(x,y)$	n-th FM component.
$a_n(x,y)\cos\varphi_n(x,y)$	n-th AM-FM component.
AM-FM Representations Based on Curvilinear Coord. Systems $(x,y)\rightarrow\Phi(x,y)=(\phi_{co1}(x,y),\phi_{co2}(x,y))$ denotes a coord. transf.	
$\exp\left(j\frac{2\pi}{T_{\text{hor}}}n\,\phi_{co1}(x,y)\right)$ $\cdot\exp\left(j\frac{2\pi}{T_{\text{vert}}}m\,\phi_{co2}(x,y)\right)$	FM harmonics derived from Φ applied to 2D periodic image with periods $T_{\text{hor}}, T_{\text{vert}}$.
$\cos\left(n\varphi_{co1}(x,y)\right)$	FM harmonic along a curv. coord.
$\exp\big(j2\pi\big(f_1\phi_{co1}(x,y)$ $+f_2\phi_{co2}(x,y)\big)\big)$	FM harmonics used with AM-FM Transforms.
Discrete-Space AM-FM Representations used with Algorithms **Symbol**	**Definition**
$I[k]$	1D defined over discrete coord. k.
$I\left[k_1,k_2\right]$	Image over discrete coords. k_1-k_2.
$\hat{I}[f_1,f_2]$	2D FFT of $I[k_1,k_2]$.
$\check{I}\left[m_1,m_2\right]$	Discrete-space FM Transform.
$I_{AS}\left[k_1,k_2\right]$	Analytic image computed using the FFT.
$I_{AS,i}\left[k_1,k_2\right]$	Analytic image processed through the i-th channel filter.
$a_n[k_1,k_2],\varphi_n[k_1,k_2]$	Discrete-space of IA and IP components.
$a_{c,i},\phi_{c,i},\nabla\phi_{c,i}$	c-channel estimates of IA, IP, and IF.
$a_n[k]\cos\phi_n[k]$	1D AM-FM comp. assoc. with $\texttt{IMF}_n[k]$.
$\exp\big(j\frac{2\pi}{N}(m_1\phi_1[k_1,k_2]$ $+m_2\phi_2[k_1,k_2])\big)$	Discrete-space FM harmonics.

is to apply a coordinate transformation to an ideal, periodic image I_{ideal} with periods T_{hor} and T_{vert} to get an AM-FM image: $I(x,y)=a(x,y)\cdot I_{\text{ideal}}\left(\phi_{co1}(x,y),\phi_{co2}(x,y)\right)$, where ϕ_{co1}, ϕ_{co2} denote the curvilinear coordinates associated with the coordinate transformation. Then, the original Fourier-Series becomes a 2D AM-FM series as given by:

$$I(x,y)=\sum_n\sum_m a_{n,m}\cdot\exp\left(j\frac{2\pi}{T_{\text{hor}}}n\,\phi_{co1}(x,y)\right)\cdot\exp\left(j\frac{2\pi}{T_{\text{vert}}}m\,\phi_{co2}(x,y)\right). \quad (2)$$

In [6], the authors describe how to use AM-FM demodulation over a Gabor filterbank to isolate the fundamental AM-FM harmonics. We shall describe general methods that extend this approach. The approach can also be simplified for the case when periodicity only extends in a single curvilinear coordinate. For example, as shown in [11] and [8] for fingerprints and woodgrain images, (2) reduces to the FM harmonics expression:

$$I(x,y)=\sum_n a_n(x,y)\cos\left(n\varphi_{co1}(x,y)\right). \quad (3)$$

On the other hand, using the same transformation on the Fourier Transform, as described in [12] and [13], we can derive a continuous-space AM-FM representation given by:

$$I(x,y)=a(x,y)\iint_{\mathbb{R}^2}\check{I}(f_1,f_2)\cdot\exp-\Big(j2\pi\big(f_1\phi_{co1}(x,y)+f_2\phi_{co2}(x,y)\big)\Big)\,df_1df_2 \quad (4)$$

where $\check{I}$ denotes the AM-FM transform, $a(x,y)>0$, and $\check{I}$ can be computed over the image plane using:

$$\check{I}(f_1,f_2)=\iint_{\mathbb{R}^2}\frac{I(x,y)}{a(x,y)}\exp\Big(j2\pi\big(f_1\phi_{co1}(x,y)+f_2\phi_{co2}(x,y)\big)\Big)\left|\det\left(\text{jac}(\Phi(x,y))\right)\right|dx\,dy \quad (5)$$

where $\Phi(\mathbb{R}^2)$ refers to the mapping of $\mathbb{R}^2$ by Φ to the region defined by (ϕ_{co1},ϕ_{co2}) and where $\det\left(\text{jac}(\Phi(x,y))\right)$ denotes the determinant of the Jacobian of the coordinate transformation. The AM-FM transform expressions of (4), (5) provide a general framework for expanding arbitrary input images.

In [9], the authors demonstrated a general method for deriving effective curvilinear coordinate systems that describe non-stationary image content as a multidimensional frequency modulation process. The approach is based on defining a coordinate system based on the local eigenvectors of the instantaneous frequency tensor.

B. Orthogonal FM Transforms as generalizations of the Discrete Fourier Transform and the Discrete Cosine Transform

Over square images $N\times N$ (where $M=N$), using integer coordinates k_1-k_2, we have the development of orthogonal FM transforms given by [14]:

$$I\left[k_1,k_2\right]=\frac{1}{N}\sum_{m_1=0}^{N-1}\sum_{m_2=0}^{N-1}\check{I}\left[m_1,m_2\right]\cdot\exp\left(j\frac{2\pi}{N}\left(m_1\phi_1[k_1,k_2]+m_2\phi_2[k_1,k_2]\right)\right) \quad (6)$$

where the discrete-space FM transform is given by

$$\check{I}\left[m_1,m_2\right]=\frac{1}{N}\sum_{k_1=0}^{N-1}\sum_{k_2=0}^{N-1}I\left[k_1,k_2\right]\cdot\exp\left(-j\frac{2\pi}{N}\left(m_1\phi_1[k_1,k_2]+m_2\phi_2[k_1,k_2]\right)\right). \quad (7)$$

The authors showed that orthogonal FM transforms as described by (6) and (7) are generalizations of the Discrete Fourier Transform (DFT) and the Discrete Cosine Transform (DCT). In particular, the basic solution is to have the forward FM transform consist of a permutation of the input 2D samples followed by a DFT (or DCT), while the inverse FM transform would then be an inverse DFT (or DCT) followed by the inverse permutation. They then showed how to compute optimal permutations with ideal FM spectra based on the permutation optimization results given in [15].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

C. Multiscale AM-FM Representations

Multiscale AM-FM representations aim to extract AM-FM components from different scales. The scales are defined in terms of the instantaneous frequency magnitude as discussed in the introduction. Here, we note that AM-FM components are estimated using discrete-space bandpass filters and thus cannot be accurately estimated near DC. Based on this observation, in [16], the authors introduced multiscale AM-FM representations as given by:

$$I\left[k_1,k_2\right] \approx d + c_0\, G[k_1,k_2] + \sum_{n=1}^{s} c_n\, a_n[k_1,k_2]\cos\varphi_n[k_1,k_2]. \quad (8)$$

where d denotes the DC component, $G[k_1,k_2]$ denotes the output of a low-pass filter, the AM-FM components: $a_n[k_1,k_2]\cos\varphi_n[k_1,k_2]$ are computed over different scales, and the coefficients $c_0, c_1, \ldots, c_s$ are computed so as to minimize the least-squares error in the reconstruction. An advantage of multiscale decompositions is that they allow us to define scales for specific applications. The introduction of multiscale decompositions with robust AM-FM demodulation methods described in [16] has led to several applications in medical image analysis.

III. AM-FM Methods

We return to Fig. 1 to introduce the basic AM-FM methods. After preprocessing, we can compute an analytic extension of the image and then apply a filterbank to produce channel-based AM-FM components. To estimate individual AM-FM components, we need to process the channel-based estimates. We then generate the AM-FM features that are used for classification. We summarize the notation used throughout the paper in Table I. While it is important to note that different authors use very different notations, we have also tried to explain how their notations relate to our standardized notation of Table I.

A. Analytic Image

The analytic image extension can be computed effectively through the application of the 1D Hilbert-transform along each row as given by [17]:

$$I_{AS}\left[k_1,k_2\right] = I\left[k_1,k_2\right] + j\mathcal{H}_{1D}\left\{I\left[k_1,k_2\right]\right\} \quad (9)$$

$$\approx \sum_{n=1}^{K} a_n[k_1,k_2]\exp\left(j\varphi_n[k_1,k_2]\right). \quad (10)$$

Compared to the original AM-FM model of (1), in (10), we can see that the real-valued FM components: $\cos\varphi_n[k_1,k_2]$ are replaced by their exponential forms: $\exp\left(j\varphi_n[k_1,k_2]\right)$. Furthermore, the computation of the analytic signal is very straightforward. To implement (9), we simply take the 2D FFT of the input image $I[k_1,k_2]$, zero-out the left two-quadrants with negative horizontal frequency component frequencies, keep the line of zero horizontal frequency components unchanged, double the amplitudes of the remaining frequencies, and take the inverse 2D FFT. Based on this approach, it is easy to see that the approximation in (10) holds exactly for the 2D FFT harmonics. In other words, for images of size $N\times M$, it is easy to see that $A\cos\left[2\pi n_1\, k_1/N + 2\pi m_1\, k_2/M\right]$ gets transformed to $A\exp\left[j2\pi n_1\, k_1/N + j2\pi m_1\, k_2/M\right]$ (except for the line of zero horizontal frequency components).

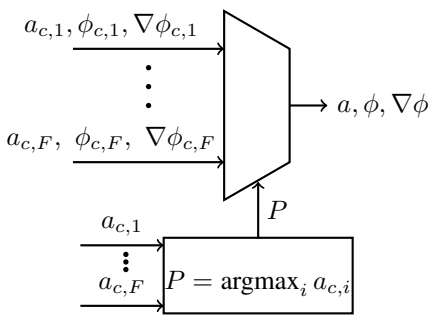

Fig. 3. Dominant Component Analysis over different scales. In this example, the c is used to index the different scales. Within each scale, at each pixel, the largest IA estimate is used to select the IA, IF, and phase from the dominant filter.

In [18], the authors provide a review of the literature on multidimensional Hilbert transform extensions and define the Hypercomplex Signals based on the quarterionic Fourier transform. Alternatively, the monogenic signal is based on replacing the Hilbert transform by the Riesz transform as described in [19], [20].

B. A Baseline Approach Based on Multiscale Filterbanks

After computing the analytic image, we should use a filterbank to reject out-of-band noise and to isolate the components. Our assumption here is that the filterbank contains a number of filters that clearly exceeds K, the number of AM-FM components given in (10). In what follows, we present a framework for using filterbanks for estimating AM-FM components. In subsequent sections, we will also provide variations of this framework.

Let $h_{c,i}$ denote the impulse responses of the filterbank associated with scale c. For each filter, we generate channel-based AM-FM components as given by:

$$I_{AS,i}\left[k_1,k_2\right] = I_{AS}\left[k_1,k_2\right] * h_{c,i}[k_1,k_2] = a_{c,i}[k_1,k_2]\exp\left(j\varphi_{c,i}[k_1,k_2]\right), \quad (11)$$

where $*$ denotes the convolution operator. We can then estimate the AM-FM components using channel component analysis [17], [21]:

$$a_{c,i}[k_1,k_2] = \frac{\left|I_{AS,i}\left[k_1,k_2\right]\right|}{H_{c,i}(\nabla\varphi_{c,i}[k_1,k_2])} \quad (12)$$

$$\varphi_{c,i}[k_1,k_2] = \arctan\left(\frac{\text{imag}(I_{AS,i}\left[k_1,k_2\right])}{\text{real}(I_{AS,i}\left[k_1,k_2\right])}\right), \quad \text{and} \quad (13)$$

$$\nabla\varphi_{c,i}[k_1,k_2] = \text{real}\left(-j\frac{\nabla I_{AS,i}\left[k_1,k_2\right]}{I_{AS,i}\left[k_1,k_2\right]}\right), \quad (14)$$

where $H_i(\nabla\varphi_{c,i}[k_1,k_2])$ denotes the amplitude magnitude of the ith filter at the estimated instantaneous frequency $\varphi_{c,i}[k_1,k_2]$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

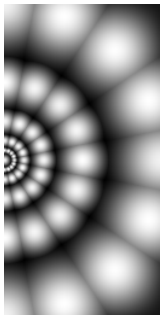

Fig. 4. Gabor Filterbank with 8 orientations and 5 scales. To visualize the filterbank, filters intersect at a quarter of the peak value in the example. Typically, filters will intersect at half-peak

Next, the goal is to use the channel-based AM-FM components to estimate the original AM-FM components. In *Dominant Component Analysis* [22], the approach is to adaptively select the channelized estimate with the largest instantaneous amplitude at each pixel. More generally, in scale-based methods, the approach was modified to restrict the selection process to channels that share a specific range of instantaneous-frequency magnitudes (e.g., low, medium, and high).

Dominant component analysis (DCA) applied over different scales is demonstrated in Fig. 3. In Fig. 3, based on the selected scale, we generate AM-FM channel estimates:

$$a_{c,i}, \phi_{c,i}, \nabla\phi_{c,i}, \quad i = 1, \ldots, F.$$

At every pixel, we select the filter that gives the maximum instantaneous amplitude. In Fig. 3, the dominant IA is given by the P-th channel. We then select $a_{c,P}, \phi_{c,P}, \nabla\phi_{c,P}$ as the estimates associated with the dominant scale component. The procedure is repeated for every pixel in the image. The single-scale AM-FM component estimate is then formed from the collection of dominant estimates over the entire image.

Filterbank design can be varied depending on the application. We summarize the most popular designs into:

- ***Gabor Filterbanks*** provide excellent localization in the spatial-frequency plane and have been very popular in computer vision applications [23]–[26]. An example with 8 orientations and 5 scales is given in Fig. 4.
- ***Separable Multiscale Filterbanks*** are designed to have flat passbands to eliminate the need the correction in (12). Thus, the channel component estimates are accurately estimated using: $a_{c,i}[k_1, k_2] = |I_{AS,i}[k_1, k_2]|$ (see [16]).
- ***Steerable Filterbanks*** were used in [27] based on [28], [29]. Steerable pyramids decompose images into scale and orientation selective subbands based on polar-separable directional derivative operators.

C. Robust estimation

To support robust estimation, we introduce three methodological improvements recommended in [16]: (i) optimal filterbanks for estimating the IA, (ii) post-filtering the maximum filter indexes in Fig. 3, and (iii) robust IF estimation using two samples. The combined use of these approaches result in better estimates of the IA, the IF, and better and continuous AM-FM component estimates overall.

We begin with the recommendation for optimal filterbanks. The basic idea is to use separable 1D filterbank design to achieve unit gain over the passband in each filter. As a result, the IA estimation formula of (12) is simplified by having $H_{c,i}(\nabla\varphi_{c,i}[k_1, k_2]) \approx 1$ to become $a_{c,i}[k_1, k_2] = |I_{AS,i}[k_1, k_2]|$. The advantage of the approach is that the variations of $H_{c,i}(\nabla\varphi_{c,i}[k_1, k_2])$ will no longer affect the estimation of the IA.

To support a better estimation of the original AM-FM components $a_n(x, y)\cos\varphi_n(x, y)$ given in (1), we need to support spatial continuity of the AM-FM components. The advantage of spatial continuity was shown in [9], where the continuity of the IF estimates was used to derive and solve the differential equations that describe the frequency modulation process throughout the image. In [16], spatial continuity was facilitated by post-filtering the maximum IA filters indexes P shown in Fig. 3. In [16], the P image, representing the filter index that gave the largest IA value at each pixel, was post-processed using a median filter to avoid oscillating between filters from pixel to pixel. As a result, severe artifacts and discontinuities were removed from the resulting AM-FM estimates.

Third, in [16], the authors developed a robust method for estimating the IF from two samples. To support the idea, prior to computing the analytic signal, for each central frequency of each filter $(\omega_{i,i}, \omega_{2,i})$, we can apply heterodyning using:

$$I[k_1, k_2] \cdot \exp(j\omega_{i,i}k_1 + j\omega_{2,i}k_2) \tag{15}$$

to slow-down high-frequency content. After this step, we put the result through each filter and apply the analytic signal extension as before. If heterodyning is applied, we will need to correct the final instantaneous frequency estimates as we describe below. The proposed robust-demodulation approach serves as an alternative to (14). To introduce the approach, we will consider the estimation of $\partial\hat{\varphi}_{c,i}[k_1, k_2]/\partial k_1$. The estimation of $\partial\hat{\varphi}_{c,i}[k_!, k_2]/\partial k_2$ is very similar.

Let the sum of two analytic output filter outputs be given by:

$$S[n_1] = I_{AS,i}[k_1 + n_1, k_2] + I_{AS,i}[k_1 - n_1, k_2]. \tag{16}$$

where the integer spacing between the samples, $2n_1$ is allowed to vary. To generate a robust IF estimate with minimum condition number, we look for $n_{1,opt}$ that minimizes the absolute sum

$$\begin{aligned} &\min_{n_1} \ |S[n_1]| \\ \text{subject to:} \quad &2n_1 \cdot (\partial\varphi_{c,i}[k_1, k_2]/\partial k_1) \leq 2\pi, \quad \text{and} \\ &\partial\varphi_{c,i}[k_1, k_2]/\partial k_1 \in \text{SR}_{1,i}. \end{aligned} \tag{17}$$

In terms of the constraints, $2n_1 \cdot (\partial\varphi_{c,i}[k_1, k_2]/\partial k_1) \leq 2\pi$ restricts the distance between the samples to be within a single period, and $\partial\varphi_{c,i}[k_1, k_2]/\partial k_1 \in \text{SR}_{1,i}$ requires that the estimated instantaneous frequency does not fall outside the first-component spectral support of the i-th filter.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

After solving (17), we plug the optimal value $n_{1,opt}$ into (16) to get the IF x-component estimate using:

$$\varphi_x[k_1,k_2] \approx \frac{1}{n_1}\arccos\left(\frac{S[n_{1,opt}]}{2I_{AS,i}\left[k_1,k_2\right]}\right) \tag{18}$$

In (18), if heterodyning pre-processing was applied, then we will need to add the offset frequency $\omega_{i,i}$ to φ_x to get $\omega_{i,i}+\varphi_x$ as the final IF component estimate. Overall, the basic idea is to evaluate the $\arccos(.)$ function near zero where it has the lowest (near-zero) condition number and can thus yield very robust estimates that are not affected by small, noisy variations.

D. AM-FM Demodulation Based on Teager Kaiser Energy Operators

The use of Teager Kaiser energy operators allows us to redefine the baseline algorithm of section III-B to be based on estimated signal energy content. Firstly, instead of selecting the channel filter with the maximum IA (e.g., see Fig. 3), the channel filter is selected based on an estimate of channel energy. Secondly, instead of (12) and (14), energy operators are used to estimate the IF and IA.

We begin with the multidimensional energy operator as given by [30], [31]:

$$\Psi\left\{I(x,y)\right\} \triangleq ||\nabla I(x,y)||^2 - I(x,y)\nabla^2 I(x,y) \tag{19}$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ denotes the Laplacian operator. Under realistic assumptions [31], applying Ψ to $a(x,y)\cos\varphi(x,y)$ will yield the signal energy estimate given by:

$$\Psi\left\{a(x,y)\cos\varphi(x,y)\right\} \approx a(x,y)^2\|\nabla\varphi\|^2. \tag{20}$$

As mentioned earlier, the estimated energy given by (20) is then used to select the channel with the maximum energy response, replacing the role of the IA [32]. The IA and IF are then estimated using [31]:

$$a(x,y) = \frac{\Psi\{I(x,y)\}}{\sqrt{\Psi\{I_x(x,y)\}+\Psi\{I_y(x,y)\}}} \tag{21}$$

$$|\varphi_x(x,y)| = \sqrt{\frac{\Psi\{I_x(x,y)\}}{\Psi\{I(x,y)\}}} \tag{22}$$

$$|\varphi_y(x,y)| = \sqrt{\frac{\Psi\{I_y(x,y)\}}{\Psi\{I(x,y)\}}} \tag{23}$$

More recently, Kokkinos *et al.* [32] introduced a regularized 2D energy operator and 2D Energy Separation Algorithm in order to obtain more accurate amplitude and frequency estimations. A multidimensional higher order Teager-Kaiser operator was introduced in [33], [34], [35] Teager-Kaiser energy methods for signal and image analysis were recently reviewed by Boudraa and Salzenstein in [36].

In [37], the authors introduced 2D multirate frequency transformations to support effective demodulation of wideband images. The basic idea is to convert wideband AM-FM components to narrowband signals that can provide more accurate estimates. For the conversion, each image is first interpolated. Then the resulting image is multiplied by $\cos(\Omega_1 k_1)\cos(\Omega_2 k_2)$ to shift the 2D frequency content to be centered at (Ω_1,Ω_2). After heterodyning, a 2D bandpass filter is applied to reduce the resulting bandwidth, and to allow effective demodulation. To recover the AM-FM component, (Ω_1,Ω_2) is added to the estimated instantaneous frequency and the resulting image is decimated to the input image resolution.

E. Empirical Mode Decomposition

In the Empirical Mode Decomposition (EMD) approach, the input signals are expressed as a sum of zero-mean AM-FM components called intrinsic mode functions (IMFs) and a residue [38]. IMFs require that (i) the number of extrema be the same as the number of zero-crossings or differ by at most by one, and (ii) at any point, the mean value of the two envelopes defined by the local maxima and minima be zero. Based on these assumptions, the IMFs are computed using the sifting process algorithm as described in Fig. 5.

In 1D, the one-dimensional AM-FM decomposition is given by:

$$\begin{aligned} I[k] &\approx \sum_{n=1}^{K} a_n[k]\cos\varphi_n[k] \\ &= \sum_{n=1}^{M} \mathtt{IMF}_n[k] + R[k]. \end{aligned} \tag{24}$$

where $\mathtt{IMF}_n[k]$ is the n-th AM-FM component and $R[k]$ represents the residue of the reconstruction. The number of IMF components is determined by the stoppage criterion as described in [39]. As discussed in [39], a large repetition of the sifting process will produce a collection of FM-only components. Thus, to produce AM-FM components, the process is terminated after the sifting process produces the same number of zero-crossings and extrema after $S=3$ to $S=5$ iterations. By varying S, as discussed in [39], the authors generate confidence limits for the extracted IMFs. Another important property of the Sifting process is that is computes low-frequency components first and then proceeds to higher frequency components.

We note that the sifting process does not use bandpass filters to reject out-of-band noise. To recognize how the approach works with noise, note that [40] showed that the EMD decomposition of fractional Gaussian noise is similar to the application of dyadic filter banks associated with Wavelet decompositions.

1) The Ensemble Empirical Mode Decomposition Method: There are two fundamental problems with the original Sifting process. First, we have mode mixing issues. In mode mixing, two or more AM-FM signals get decomposed into different IMFs. We thus get estimation artifacts that appear as very disparate amplitudes in individual IMFs or very similar oscillations in different IMFs. Secondly, small perturbations can yield completely different IMFs [41]. To address these issues, the authors in [42], provided a noise assisted method, termed the Ensemble Empirical Mode Decomposition (EEMD).

We summarize the EEMD into four steps [42]. First, we add white noise to the original data. Second, we apply the Sifting method to the noise corrupted signals. Third, we repeat the first

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

function SIFTING($I[k]$)
  while not stoppingCriterion **do**
    $[\texttt{Maxima}, \texttt{Minima}] \leftarrow \texttt{extrema}(\texttt{x[k]})$
    $e_{max}[k] \leftarrow \texttt{fitCubicSplines}(\texttt{Maxima})$
    $e_{min}[k] \leftarrow \texttt{fitCubicSplines}(\texttt{Minima})$
    $m[k] = (e_{min}[k] + e_{max}[k])/2$
    $h[k] = I[k]$ - $m[k]$
    $I[k] \leftarrow h[k]$
  end while
  $\texttt{IMF} \leftarrow I[k]$
end function

Fig. 5. Sifting process the algorithm for computing the EMD.

two steps for different noise levels. Fourth, we take the means of IMF estimates as the final estimates. The approach also requires the development of new stoppage criteria as detailed in [42].

The original EEMD method was improved in the Complete Ensemble Empirical Mode Decomposition (CEEMD) [43]. For CEEMD, the first mode is the same as the one generated by EEMD as the average AM-FM component estimate. For the second mode, we initially form the residual between the original signal and the first EEMD estimate as given by:

$$r_1[k] = I[k] - \widetilde{\texttt{IMF}}_1[k].$$

Then, the second mode ($\widetilde{\texttt{IMF}}_2$) is set to the average over the first modes extracted by the EMD algorithm when applied to an ensemble of r_1 with different levels of noise. The second residual is then formed using:

$$r_2[k] = r_1[k] - \widetilde{\texttt{IMF}}_2[k].$$

From here, the process is repeated to generate the remaining modes.

2) Multi-dimensional Extensions to the Empirical Mode Decomposition: We provide a summary of EMD extensions for digital images. We classify the methods into separable approaches that were applied along the rows and columns, and 2D methods that work with 2D extrema.

An early attempt to extend EMD to two dimensions treated 2D images as a collection of 1D slides using 1D EMD along each 1D slide (see [44]). This early attempt gives preference to a particular direction and is not appropriate for use in digital images.

More recently, the multi-dimensional empirical mode decomposition (MEEMD) extended the EEMD method of section III-E1 to avoid biases in any particular direction [45]. The approach can be summarized into three steps. First, the EEMD is applied along each row. Second, the ensemble of 2D image outputs are input to the EEMD to generate ensemble outputs along each column. Third, if we let $h_{i,j}$ denote the outputs of the second step, the final AM-FM components C_i are generated using equation (13) of [45] which is:

$$a_n[k_1, k_2]\cos\varphi_n[k_1, k_2] = \sum_{k=n}^{K} h_{i,k} + \sum_{j=n+1}^{J} h_{j,i} \qquad (25)$$

where K and J denote the ensemble sizes for the columns and rows respectively. The basic idea in (25) is to generate component estimates that have comparable minimal scales along the rows and columns.

The Multi-dimensional Complete Ensemble Empirical Mode Decomposition with Adaptive noise (MCEEMDAN) extend the one-dimensional CEEMD by applying the CEEMD across the rows and then the columns. We refer to [46] for details.

We provide a brief summary of different variations of the basic ideas presented here. For the Bidimensional EMD [47] (BEMD), the authors applied new methods for estimating 2D extrema and 2D interpolation. In [47], the authors use morphological reconstruction to extract 2D minima and maxima. Then, instead of using cubic spline interpolation, the authors use radial basis functions. In [48], the authors used the BEMD with the monogenic signal (see [19], [20]) to analyze different textures.

3) Variational Mode Decomposition: To describe the variational algorithms, we begin with defining the inner-product for two functions $f(x, y)$ and $g(x, y)$ using:

$$\langle f(x,y), g(x,y)\rangle = \iint_{\mathbb{R}^2} f(x,y)g(x,y)dx\,dy.$$

Using the inner-product, we define the $\|.\|_2^2$ norm using:

$$\|f(x,y)\|_2^2 = \langle f(x,y), f(x,y)\rangle .$$

For the variational decomposition algorithms, the assumption is that AM-FM components are centered around central frequencies: $\omega_1, \omega_2, \ldots, \omega_K$, where each n-th component central frequency is two-dimensional: $\omega_n = (\omega_{n1}, \omega_{n2}) = 2\pi(f_{n1}, f_{n2})$. Following the authors, we use $\{\omega_n\}$ to describe the set of $\omega_1, \omega_2, \ldots, \omega_K$.

Since the components are assumed to be concentrated along their central frequencies, the assumption is that the baseband bandwidth is small and thus the $\|.\|_2^2$ norm of the gradient functional around each component given by:

$$\text{GradFunc}(\omega_n) = \left\| \nabla\left(I_{AS}(x,y)\exp\left(-j(\omega_{n1}x + \omega_{n2}y)\right)\right)\right\|_2^2 \qquad (26)$$

is assumed to be small. Clearly, minimizing (26) alone will not reconstruct general AM-FM components that are not captured by pure sinusoids. Thus, to avoid this issue, the method requires minimization of the reconstruction error given by:

$$\text{RecError}(\{u_n\}) = I(x,y) - \sum_{n=1}^{K} u_n(x,y),$$

where $\{u_n\}$, as before, represents the set of $u_1(x,y), \ldots, u_K(x,y)$ which are the estimated AM-FM components $u_n(x,y) = a_n(x,y)\cos\varphi_n(x,y)$. Then, the overall optimization method described in [49] (extending [50]) is given by:

$$\sum_{n=1}^{n=K} \alpha_n \text{GradFunc}(\omega_n) \qquad (27)$$

$$\text{subject to:} \quad \text{RecError}(\{u_n\}) = 0. \qquad (28)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

where the α_n coefficients are used as weights for controlling the spread around each AM-FM component. In order to solve the constraint optimization problem, we form the Lagrangian

$$\mathcal{L}\left(\{\omega_n\},\{u_n\},\rho,\lambda\right) = \sum_{n=1}^{n=K} \alpha_n \text{GradFunc}(\omega_n) + \text{EqConstr}(\rho,\lambda,\text{RecError}(\{u_n\})) \quad (29)$$

where the equality constraint is given by

$$\text{EqConstr}(\rho,\lambda,\text{RecError}(\{u_n\})) = \rho\left\|\text{RecError}(\{u_n\})\right\|_2^2 + \langle\lambda(x,y),\text{RecError}(\{u_n\})\rangle \quad (30)$$

where ρ represents a Lagrange multiplier and $\lambda(x,y)$ represents a Lagrange multiplier function needed to solve the constrained optimization problem. We solve (29) using:

$$\min_{\{\omega_n\},\{u_n\}} \max_{\rho,\lambda(x,y)} \mathcal{L}\left(\{\omega_n\},\{u_n\},\rho,\lambda\right). \quad (31)$$

In order to solve (31), we apply Parseval's theorem to transform (26) to the frequency domain. After that, we apply coordinate descent optimization over each AM-FM component u_n sequentially and use the centroid frequency of the resulting analytic signal to estimate ω_n. Once all of the AM-FM components have been estimated, the Lagrangian function is updated in the frequency domain (using the 2D FFT) using ascent on the reconstruction error:

$$\hat{\lambda}^{\texttt{iter}+1}[f_1,f_2] = \hat{\lambda}^{\texttt{iter}}[f_1,f_2] + \tau\left(\hat{I}[f_1,f_2] - \sum_{n=1}^{n=K} \hat{u}_n^{\texttt{iter}+1}[f_1,f_2]\right) \quad (32)$$

where `iter` refers to the iteration number and the rest of the parameters are set to 1 (see [49]). The algorithm iterates until there is little relative change in the estimates of $\hat{u}_n^{\texttt{iter}}$ between iterations.

The method was extended in [51] to account for spatially-bounded components. For the new model, let $\{S_n(x,y)\}$ denote binary (0 or 1) segmentation functions associated with the corresponding AM-FM components $\{u_n(x,y)\}$. The approach redefines the reconstruction error to account for the spatial limitations of each component as given by:

$$\text{RecError}\left(\{u_n\},\{S_n\}\right) = I(x,y) - \sum_{n=1}^{K} S_n(x,y)u_n(x,y). \quad (33)$$

Then, in order to estimate the n-th segmentation function, the method requires the minimization of an additional term given by:

$$\text{SegmFunc}(\beta_n,\gamma_n,S_n) = \beta_n\|S_n(x,y)\|_1 + \gamma_n\,\text{TV}(S_n) \quad (34)$$

where we define the $\|.\|_1$ norm using:

$$\|S_n(x,y)\|_1 = \iint_{\mathbb{R}^2} |S_n(x,y)|dx\,dy,$$

the total variation term is given by:

$$\text{TV}(S_n) = \sup_{\|G\|\leq 1} \left\langle S_n(x,y),\,\text{div}\left(G(x,y)\right)\right\rangle,$$

the divergence operator is given by:

$$\text{div}\left(G(x,y)\right) = \frac{G_x(x,y)}{\partial x} + \frac{G_y(x,y)}{\partial y}.$$

The modified Lagrangian is given by:

$$\mathcal{L}(\{\omega_n\},\{u_n\},\{v_n\},\{S_n\},\{\alpha_n\},\{\beta_n\},\{\gamma_n\},\rho,\lambda,\{\rho_n\},\{\lambda_n\}) = \sum_{n=1}^{K}\Bigg(\text{GradFunc}(\alpha_n,\omega_n) + \text{SegmFunc}(\beta_n,\gamma_n,S_n) + \text{EqConstr}\left(\rho_n,\lambda_n(x,y),u_n(x,y)-v_n(x,y)\right)\Bigg) + \text{EqConstr}\left(\rho,\lambda,\text{RecError}(\{u_n\},\{S_n\})\right) \quad (35)$$

where the auxiliary functions v_n and the equality constraints $u_n = v_n$ are used by the optimization algorithm to relax the constraints and ensure spectral solvability [51].

To minimize the Lagrangian of (35), the authors recommend the use of coordinate descent through the different functions as described for (31). The coordinate descent steps for $\{\omega_n\}$ $\{u_n\}$ and $\{v_n\}$ are the same as for (31). The big difference comes from the need to minimize the segmentation functions $\{S_n\}$ that use the TV norm. The minimization steps include: Repeat for $t = 1$ to T:

1. Apply area penalty and reconstruction fidelity ODE:

$$\frac{\partial S_n}{\partial t} = -\beta_k + 2\rho\,v_n(x,y) \cdot\left(\text{RecError}(\{u_n\},\{S_n\}) + \frac{\lambda(x)}{\rho}\right) \quad (36)$$

2. Apply diffusion equation to generate smoothed regions:

$$\frac{\partial S_k}{\partial t} = \gamma_k\nabla^2 S_k(x), \quad (37)$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$.

3. Generate segmented image via thresholding $S_k > 0.5$.

The three steps iterate from (1) minimizing the reconstruction error, (2) diffusion of the segmentation function to approximate piecewise constant values, and (3) a thresholding step to generate binary segmentation functions. We provide an example application and discuss computational complexity issues in section IV.

F. Machine Learning Methods

Following the computation of AM-FM decompositions, there is a need to select appropriate features for different classification purposes (refer to Fig. 1). Compared to the current emergence of Deep Learning methods, the approaches described here are simple, easy to interpret, and surprisingly effective (e.g., see [6]).

We begin with a brief summary of methods used for feature extraction and selection. At the lowest level, we have the use of histograms of the IA, IF magnitudes (|IA|, |IF|) and IF angle ($\angle$IF). In such approaches, strong features are associated with high IA. A simple threshold on the IA is often used to reject out of band noise while higher IF magnitudes are associated

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

with fine textures. For detecting and classifying regions of interest refer to [6] and multiscale examples in [7]. Similarly, in EMD analysis, we also have the use of multiple scales to describe texture. Nunes *et al.* [47], [48], [52] studied texture analysis based on BEMD using the instantaneous amplitude, phase, isotropy and orientation to describe texture variations.

Despite the large number of generated AM-FM features, AM-FM feature selection has only been done using traditional statistical tests (e.g., using the Wilcoxon Rank sum test or a t-test). Statistical tests have also been applied to specific histogram bins of the distributions of AM and FM features. Yet, the extracted IA and IF component features are of the same size as the original input images. There is a clear need for feature research in effective feature research using deep learning methods (sec VI-C).

Following feature extraction and feature selection, we perform classification as shown in Fig. 1. AM-FM classifiers have been limited to support vector machines (SVM), partial least squares, often combined with K-means clustering methods. Although implemented for 1D AM-FM signals, it is still interesting to note the application of Independent Component Analysis and SVM described in [53]. More recently, the importance of using FM images as opposed to raw images with deep learning methods has been demonstrated in [54]. Thus, there is a strong need for the development of medical image analysis applications that use AM-FM features with deep learning classifiers. We refer to section VI-C for recommendations.

IV. Experimental Comparisons

We will next provide comparisons between three different AM-FM methods: (i) multiscale AM-FM (MS-AM-FM) method using a Gabor filterbank (section III-C), (ii) the MEEMD method (see section III-E2), and (iii) the Variational Mode Decomposition (VMD) for the modified Lagrangian of (35) (see section III-E3). The Matlab code is available at https://ivpcl.unm.edu or http://www.ehealthlab.cs.ucy.ac.cy (including datasets) for MS-AM-FM, where MEEMD was taken from the Appendix of [45], and VMD from http://www.math.montana.edu/dzosso/code.

To compare among the different methods, we consider a three-component image corrupted by white Gaussian noise. For the three components, refer to Figs. 6(a)-(d). The input image contains three AM-FM components from three different scales, corrupted by white Gaussian noise as given by:

$$\begin{aligned} I[k_1,k_2] &= \text{Comp}_1[\omega_1,\sigma_1,\alpha_1,\beta_1,x_{1,0},y_{1,0},k_1,k_2] \\ &\quad + \text{Comp}_2[\omega_2,\sigma_2,\alpha_2,\beta_2,x_{2,0},y_{2,0},k_1,k_2] \\ &\quad + \text{Comp}_3[\omega_3,\sigma_3,\alpha_3,\beta_3,x_{3,0},y_{3,0},k_1,k_2] + 5\cdot n[k_1,k_2] \end{aligned} \tag{38}$$

where $n[.]$ denotes zero-mean unit-variance white Gaussian noise, and the AM-FM components use a Gaussian envelope for the IA and quadratic FM as given by the parameter summary of Table II with each component given by:

$$\begin{aligned} &\text{Comp}_i[\omega_i,\sigma_i,\alpha_i,\beta_i,x_{i,0},y_{i,0},k_1,k_2] = \\ &\quad \exp\left[-\frac{1}{2\sigma_i^2}\left(\left(\frac{k_1-x_{i,0}}{N/2}\right)^2+\left(\frac{k_2-y_{i,0}}{M/2}\right)^2\right)\right] \\ &\quad \cdot\cos\left[\alpha_i\left(\frac{k_1-x_{i,0}}{N/2}\right)^2+(k_1-x_{i,0})\,\omega_{i,1}\right. \\ &\quad \left.+\beta_i\left(\frac{k_2-y_{i,0}}{M/2}\right)^2+(k_2-y_{i,0})\,\omega_{i,2}\right]. \end{aligned} \tag{39}$$

We also provide a summary of how we run each method. For the multiscale method, we used the standard Gabor filterbank with half-peak intersections with low IA rejection based on component areas (e.g., see [16]). For MEEMD, we set the ensemble number to 100 and set the number of modes to 3. For the variational mode decomposition, the parameters were set for texture decomposition as described in section 7.1 of [51], and number of modes was set to 4. Beyond AM-FM component component comparisons, refer to Fig. 2 for IA and IF estimation, and the description of robust IA and IF estimation in section III-C.

We provide a comparison of the estimated AM-FM components among the different methods in Fig. 6. Based on careful, close-up inspection of the results, we can see significant mode mixing for MEEMD, reduced mode mixing for the variational mode mixing, and no mode mixing for the multiscale approach. Here, we note that the multiscale approach uses low-IA rejection to reject mode interference. Except for significant mode mixing in the DC component, the variational mode decomposition provided smooth estimates. The use of a multiscale Gabor filterbank performed very well for the medium-scale and high-scale components, but suffered for the lower-scales.

In terms of computational complexity, we refer to the measured times in Table III. We note that the multiscale method was the fastest, closely followed by the variational mode decomposition. MEEMD was significantly slower. More generally, the computational complexity associated with the multi-scale method is dominated by the application of the Gabor filterbank. Essentially, for the multi-scale method, computational complexity is directly proportional to the number of filters that are used, since each filter is convolved with the extended analytic image. For the MEEMD and the variational methods, computational complexity is directly proportional to the number of components. To see this, for the variational methods, note that coordinate descent is proportional to the number of AM-FM components that need to be estimated. Similarly, for MEEMD and the rest of the EMD methods, we iterate over each component.

V. Selected Applications in Medical Image Analysis

Medical imaging applications of AM-FM analysis are presented that are briefly summarized in Table I. To keep the number of considered applications manageable, we do not consider the application of 1D AM-FM methods.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

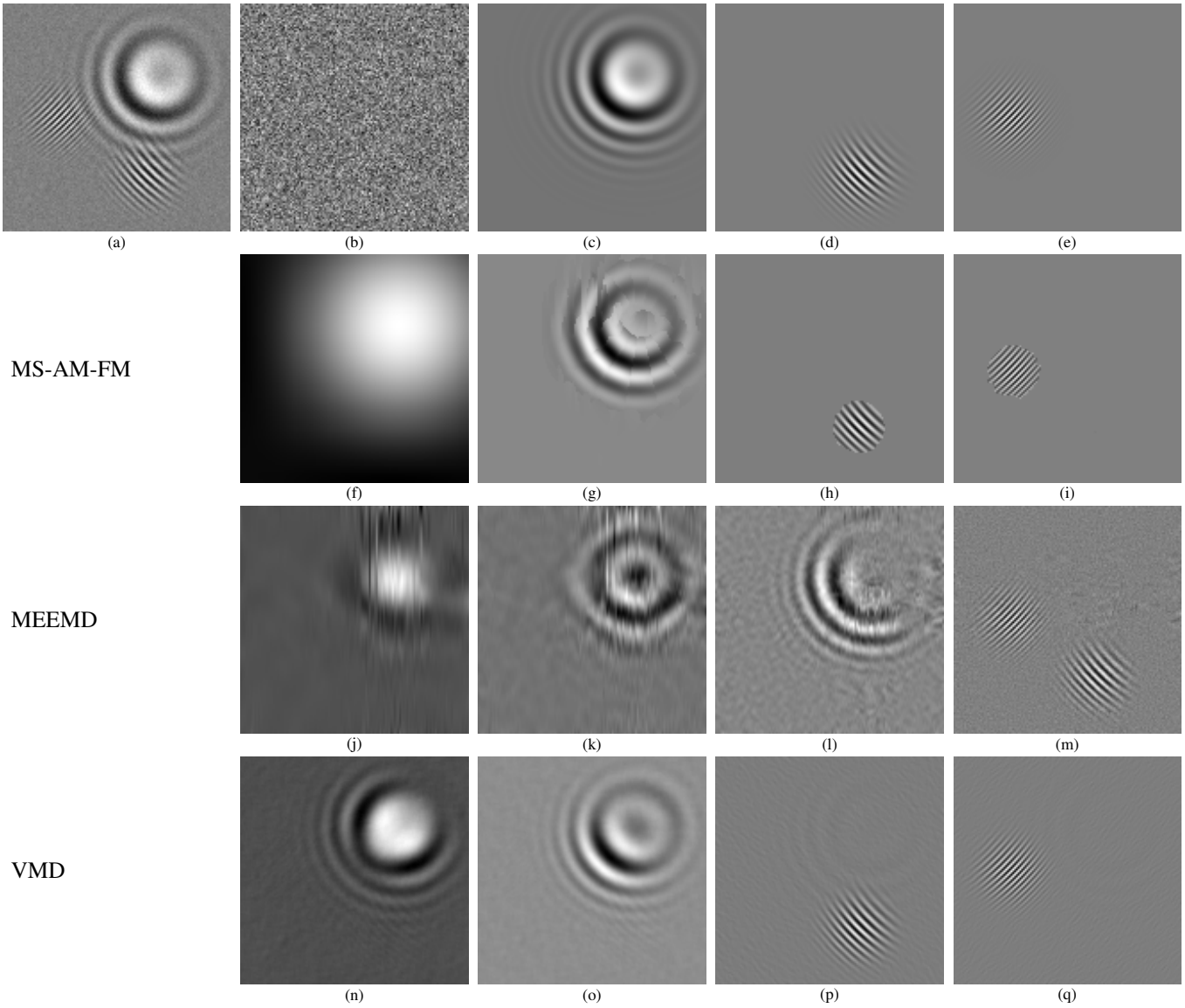

Fig. 6. A comparison of AM-FM component estimation for different methods. Ground truth: (a) synthetic image, (b) additive noise, (c) low-scale component, (d) medium-scale component. (e) high-scale component. MS-AM-FM results: (f) low-pass filter output, (g) low-scale component estimate with IA above the median value, (h) medium-scale component estimate with IA above the 96-th percentile, (i) high-scale component estimate with IA above the 96-th percentile. MEEMD results: (j) residual, (k) IMF3, (l) IMF2, (m) IMF1. VMD results: (n) DC, (o) 4th component estimate (for low-scale), (p) 3rd component estimate (for medium-scale), (q) 2nd component estimate (for high-scale). Zoom in to visualize reconstruction noise and mode mixing.

TABLE II
PARAMETERS USED FOR SYNTHETIC AM-FM COMPONENT EXAMPLE FOR SIMULATION. THE INPUT IMAGE IS 128×128 WITH THREE COMPONENTS AS GIVEN BY (38). THE EQUATION EXPLAINING HOW THE PARAMETERS ARE USED BY EACH COMPONENT ARE GIVEN IN (39).

Component	$\omega_{i,1}$	$\omega_{i,2}$	σ_i	$x_{i,0}$	$y_{i,0}$	α_i	β_i
$i=1$	$-\frac{\pi\sqrt{2}}{32}$	$\frac{\pi\sqrt{2}}{32}$	$\frac{8}{9\pi}$	$\frac{5N}{8}$	$\frac{3M}{8}$	$\frac{N}{4}$	$\frac{M}{4}$
$i=2$	$-\frac{\pi\sqrt{2}}{4}$	$\frac{\pi\sqrt{2}}{4}$	$\frac{1}{2\pi}$	$\frac{5N}{8}$	$\frac{3M}{4}$	$\frac{N}{4}$	$\frac{M}{4}$
$i=3$	$\frac{3\pi\sqrt{2}}{8}$	$\frac{3\pi\sqrt{2}}{8}$	$\frac{1}{2\pi}$	$\frac{N}{4}$	$\frac{M}{2}$	$\frac{N}{4}$	$\frac{M}{4}$

A. Cardiovascular

1) Texture Analysis of the Intima and Media Layers of the Common Carotid Artery: Early signs of atherosclerosis are associated with texture changes of the different layers of the carotid artery. Thus, the goal is to characterize texture changes prior to the formation of atherosclerotic plaques.

In [55], the authors studied the use of AM-FM features to analyze texture changes of the intima media complex, the

TABLE III
COMPUTATIONAL TIMES FOR EXTRACTING THE AM-FM COMPONENTS FOR THE EXAMPLE DESCRIBED IN TABLE II. THE EXAMPLES WERE IMPLEMENTED USING MATLAB VERSION 9.5.0 (R2018B) ON A MACBOOK PRO WITH A 2.8 GHZ INTEL I7-2640M PROCESSOR (LATE 2011 MODEL) AND 16GB RAM.

Method	Time (s)
MS-AM-FM	2.59
MEEMD	33.86
VMD	4.15

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

TABLE IV
APPLICATIONS OF AM-FM MODELS IN MEDICAL IMAGING

Reference	AM-FM Method	Dataset	Features	Results			
				AUC	Sen	Spe	Acc
Cardiovascular							
Texture Analysis of Common Carotid Artery [55]	MSAIm, DF	100 Im	IA,|IF|,∠IF Hist	NA	NA	NA	NA
Atherosclerotic Carotid Plaque Classification [56]	BEMD	1353 S, 420 AS	Entropy IMFs	NA	97%	83%	91%
Microvascular Blood Flow [46]	MCEEMDAN	30 Im	NA	NA	NA	NA	NA
Myocardial Infraction Quantification [57]	MEEMD, MS	82 P	Phase	0.754	NA	NA	NA
Nervous System							
Brain							
Multiple Sclerosis Texture Analysis [58]	MSAIm, DF	38 P 20 HC	IA,|IF|,∠IF	0.76	71%	95%	86%
Multiple Sclerosis Prediction [59]	MSAIm	20 P	Morph, Stat	NA	NA	NA	75%–97%
Alzheimer's Disease PET Image Analysis [60]	MEEMD	100 P, 100 HC	IMFs	NA	NA	NA	95%
Parkinson Disease DaTSCAN SPECT Image Analysis [61]	MEEMD	41 P 39 HC	IMFs	0.95	95%	94%	95%
Special Senses (Eye)							
Diabetic Retinopathy Lesion Detection [7]	MSAIm, SF	108 P, 70 HC	IA,|IF|, ∠IF Hist	0.84	92%	54%	92%
Diabetic Retinopathy Lesion Detection [62]	MSAIm, SF	437 AN, 136 N	IA,|IF|, ∠IF Hist	0.89	97%	50%	NA
Age-Related Macula Degeneration Lesion Detection [62]	MSAIm, SF	248 AN, 144 N	IA,|IF|, ∠IF Hist	0.84	94%	50%	NA
Macula Exudates Detection [63]	MDAIm, SF	1052 Im	IA	0.96	73%	99%	92%
Optic Disk Neovascularization Classification [64]	MDAIm, SF	300 Im	IA, |IF|, ∠IF	0.93	78%	94%	88%
Retinal Image Hemorrhage Detection [65]	VMD	108 Im	Stat	NA	100%	100%	100%
Glaucoma Detection [66]	VMD	244 Im, 244 P	Entropy, FD	NA	94%	97%	95%
Musculoskeletal							
Electron Microscopy Image Segmentation [6]	AIm, GF	26 Im	IA, IF	NA	NA	NA	84%
Respiration							
Pneumoconiosis [67]	MSAIm	22 Im	IA,|IF| Hist	1	NA	NA	NA
Gynecological Cancer							
Mammographic Density Classification [68]	MDAIm, GF	[69] Database	IA Hist	NA	NA	NA	80%
Breast Architectural Distortion Detection [70]	BEMD	187 P, 2191 HC	fractals	0.95	NA	NA	NA
Classification of Histopathological Breast Cancer Images [71]	VMD	5429 M, 2480 B	Zernike moments	NA	87%	85%	87%
Endometrium Hysteroscopy Imaging [72]	MSAIm	50 AN, 50 N	IA,IF,IF	NA	75%	90%	83%
Abdomen							
Focal Liver Ultrasonic Lesion Classification [73]	BEMD	78 N,26 B,36 M	DCT of IMFs	NA	91%	97%	93%
Hepatic Fibrosis Microscopic Classification [74]	BEMD	15 N, 45 AN	IMFs, Stat	NA	NA	NA	98%

DF:Dyadic Filterbank, GF: Gabor Filterbank, SF: Separable Filterbank, EO: Energy Operator, TKEO: Teager Kaiser Energy Operator, AIm: Analytic Image, MSAIm: Multiscale Analytic Image, MS: Monogenic Signal, ∠IF: IF Angle, HWF: Hilbert Weighted Frequency, Stat: Statistical, Morph: Morphological, Hist: Histograms,FD: Fractals, Im: Images, P: Patients, HC: Healthy Control,S: Symptomatic,AS: Asymptomatic, AN: Abnormal, N: Normal, B: Benign, M: Malignant, AUC: Area Under the Curve,

media layer and the intima layer of the common carotid artery. Here, the authors used multiscale AM-FM analysis and extracted histograms of the IA, IP, and IF for the low, medium, and high scales from each ROI (e.g., medial layer, etc). Based on 100 B-mode ultrasound images, for features extracted from the <50 and >60 age groups, the study found significant texture differences: (i) in the low-scale and medium-scale IA and high-scale IF extracted from the intima media complex, and (ii) the low-scale IA extracted from the intima layer. Furthermore, for the IA extracted from the media layer, the study found significant differences among men and women for all age groups (<50, 50-60, >60). The study showed that AM-FM features provided complimentary information that goes beyond traditional texture features such as the gray-scale median, contrast, and coarseness.

2) Atherosclerotic Carotid Plaque Classification: In [56], the authors presented an early detection framework for predicting carotid plaque deposition. The authors used the BEMD to extract entropy based features from the IMFs from a total of 1353 symptomatic and 420 asymptomatic carotid plaque ultrasound images. To improve classification accuracy and reduce bias, the dataset was balanced with adaptive synthetic sampling. An SVM classifier trained on 14 selected features gave a classification accuracy of 91%, sensitivity of 97%, and specificity of 83.2%.

3) Microvascular Blood Flow: Microvascular blood flow alternations were identified in several disease processes where numerous studies tried to evaluate the clinical implications of these changes [75]. Laser Speckle Contrast Imaging (LSCI) is a microvascular blood flow modality based on the phenomenon of laser light backscattering that forms a random interference pattern (the speckle pattern). Humeau *et al.* [76] [46] introduced the analysis of LSCI using EMD and its variations to obtain the IMFs. A comparative study between Laser Speckle

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

Contrast Imaging and Laser Doppler Flowmetry was presented in [76], using EEMD, and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise. Furthermore, their work was extended to MCEEMDAN [46] that resulted in the extraction of different patterns in different situations that need to be further evaluated.

4) Myocardial Infarction Quantification: High risk cardiovascular mortality patients can be identified prior to myocardial infraction by Gadolinium Enhancement cardiovascular magnetic resonance imaging. In [57], the authors presented a framework for myocardial infarction quantification consisting of image decomposition into intrinsic modes using multidimensional Ensemble Empirical Mode Decomposition, followed by the extraction of monogenic phase on the combined dominant modes. On the extracted phase, the authors used multilevel Otsu thresholding to quantify image regions into healthy, gray zone, and infarct core regions of the myocardium. Using a database of 82 patients, the authors obtained an AUC of 0.754 for total infarct, an AUC of 0.709 for infarct core, and an AUC of 0.845 for gray zone.

B. Nervous System

1) Multiple Sclerosis (MS): MS is a disease of the central nervous system that disrupts the flow of information within the brain and between the brain and the body that often develops disability. At this time, there are no symptoms, physical findings or laboratory tests that can, by themselves, determine if a person has MS. Magnetic Resonance Imaging (MRI) is used to provide information about disease burden and lesion evolution. The difficulty appears into white matter areas that could be possibly classified into the following three categories: Normal White Matter, Normal Appearing White Matter and MS lesions. The discrimination between these three tissue types cannot be based on their appearances. Various image analysis techniques were used to quantify MRI abnormalities in order to capture diagnostically significant image features for disease activity monitoring and progression. The Expanded Disability Status Scale (EDSS) is a clinical index scaling for the assessment of disease progression. In [58] Loizou *et al.* used multiscale AM-FM representation with Hilbert based demodulation with components extracted by a dyadic filterbank. Analysis was based on a multiscale filterbank. Low, medium and high frequency scale based feature combinations were investigated as follows: 1) IA; 2) |IF|; 3) IA; |IF|; 4) |IF| $\angle$ IF; and 5) IA |IF| $\angle$ IF.

T2-weighted transverse MR images were recorded (with repeated scans with an interval of 6-12 months) from 38 patients and 20 healthy volunteers. Disability was quantified after 2 years of the initial diagnosis as follows: 23 early cases (EDSS $\leq$ 2) and 15 advance cases (EDSS $>$ 2) of the disease respectively. Significant statistical differences were identified in the AM-FM features between NWM, NAWM and lesions utilizing the Mann-Whitney test. Furthermore, patients were classified using the SVM classifier into two classes (those with EDSS $\leq$ 2 and those with EDSS $>$ 2) achieving a correct classifications score of 86%. The results indicated that AM-FM features provided useful information for lesion load characterization and disability progression in MS.

MS progression detection was proposed by Washimkar and Chede [59]. They proposed brain MRI AM-FM segmentation followed by saliency map mapping and fuzzy c means clustering. Then, an adaptive iterative threshold based algorithm for lesion detection was applied. Morphological, local binary pattern and statistical lesion features were extracted, which were classified with the K nearest neighbor (K-NN) classifier. The accuracy of the classifier for disease detection for the first week of occurrence gave the best results and ranged from 75 to 97%.

2) Alzheimer's Disease (AD): AD is a brain disorder that in most people appears in the mid-60s, characterized of irreversible, progressive loss of memory and thinking skills. Early diagnosis is important for patient medical treatment to slow or delay the symptoms of disease. Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) are commonly used in AD diagnosis.

An application based on EMD for PET image analysis in the classification of normal, mild cognitive impairment (MCI) or AD subjects was proposed by Neubauer *et al.* [60]. In this work, normalized data were processed using MEEMD where single or a combination of Bidimensional Intrinsic Mode Functions were used as feature sets for the classification with SVM or Random Forest decision trees. Functional PET images of 100 AD or MCI subjects and 100 normal control (NC) subjects, were used for the classification of subjects suffering from AD or MCI against the NC group. 100 AD/MCI and 100 NC images, each of size 79x95 pixels, were used for the classification of subjects suffering from AD or MCI against a NC group with authors asserting classification rates in the range 90% to 95%.

3) Parkinsonian Syndrome (PS): Radiopharmaceutical (Ioflupane) brain SPECT imaging is used for PS diagnosis. Rojas *et al.* [61] provided an EMD-based approach for PS classification based on SPECT images. The method was evaluated using 80 images, yielding up to 95% accuracy.

C. Special Senses (Eye)

1) Diabetic retinopathy: Diabetic retinopathy (DR) affects blood vessels in the retina and is the most common cause of vision loss among people with diabetes. Changes to the blood vessels include microaneurysms, exudates, hemorrhages and neovascularization at different levels or grades of the disease. We will next provide a summary of the application of multiscale AM-FM methods to diabetic retinopathy [7], [62]–[64].

In [7], the authors showed that AM-FM features can be used to differentiate among microaneurysms, exudates, neovascularization on the retina, hemorrhages, normal retinal background, and normal vessels patterns (also see earlier work in [77]). The method include the use of robust estimation methods detailed in section III-C. For the specific application, the authors considered four scales implemented using a separable filterbank and nine combinations of scales for feature extraction. Feature extraction consisted of 27 cumulative distribution functions (CDFs), three for each of the nine CoS for the IA, |IF| and $\angle$IF. As shown in the paper, the CDFs exhibited strong

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

differences between different pairs of abnormal structures. For classification over the MESSIDOR database, the authors reported classification accuracies of 97% for Risk 3 (70 test images), 82% for Risk 2 (18 test images), 89% for Risk 1 (30 test images), overall accuracy of 92%. Furthermore, the paper reported an AUC of 0.84 with the recommended sensitivity/specificity pair of 92%/ 54% for Risks 3, 2, and 1 versus Risk 0.

In [63], Agurto *et al.* presented a multi-scale approach for detecting exudates regions in the macula region of the retina. The approach is based on thresholding IA components from multiple scales to generate candidate exudate regions. From each candidate region, color, shape and texture features are extracted and are fed to a partial least squares classification method. The approach yielded an AUC of 0.96 for the combination of two datasets of 400 and 652 images.

In [64], the authors developed a new method to detect abnormal vasculature in the optic disk. The approach uses an adaptive segmentation approach to select candidate vasculature regions. From each candidate region, the method extracts multiscale AM-FM features from twelve combinations of frequency bands, morphological granulometry and fractal dimension features. To test the approach, ROIs over the segmented vessels were classified as normal and abnormal. For 300 images (100 neovascularization and 200 normal), a linear SVN classifier gave a classification accuracy of 88%.

Lahmiri and Shmuel [65] proposed an automated system to detect hemorrhages in retinal images based on VMD, texture statistical descriptors, and machine learning classification. Initially, the retinal image is decomposed into a set of variational modes. Subsequently, a set of four statistical texture descriptors from the first variational mode, that represents the high frequency components, are used as features for five different classifiers in order to characterize retinal texture. For hemorrhage detection, the best classification performance of 100% accuracy was obtained with and SVM classifier.

2) Age-Related Macula Degeneration: Age-Related Macula Degeneration (AMD) refers to the decline of macula functions that affect the central vision system. AMD is the leading cause of significant visual acuity loss in people over the age of 50 in developed countries. The classification of pathological structures such as drusen (AMD characteristic deposits) is crucial for early diagnosis.

In [62], the authors described an AM-FM based system for detecting DR and AMD in digital fundus images (also see earlier work in [78]). The approach was based on the use of histograms of IA,$|$IF$|$ and $\angle$IF and statistical moments as input features to a K-means clustering method followed by the use of a partial least squares classifier. The method was tested on 2247 digital photographs from 822 patients and reported an AUC of 0.89 for detecting DR with sensitivity/specificity of 97%/50% and an AUC of 0.84 for AMD with sensitivity/specificity of 94%/50% respectively.

3) Glaucoma detection: Glaucoma is a complex disease in which damage to the optic nerve leads to progressive, irreversible vision loss. An automated glaucoma diagnosis system presented by Maheshwari *et al.* [66] using Variational Mode Decomposition. Various entropy and fractal features were extracted in order to classify 244 normal and 244 glaucoma color fundus images. The least squares support vector machine was used for classification and the best accuracy achieved was at 95% with sensitivity/specificity of 94%/97%.

D. Musculoskeletal

Myopathies such as congenital, vacuolar, and metabolic often require the visualization of the extent of the disease via the study of electron microscopy muscle biopsy images. In the case of myopathy, abnormal structures disturb the very regular and repetitive sarcomere pattern of myofibers. Pattichis *et al.* [6] proposed an AM-FM based electron microscopy image segmentation system of skeletal muscle for the recognition of normal and abnormal regions. Analysis was based on the analytic image demodulation using Gabor filterbanks and DCA. The IA and IF estimates were used for driving the segmentation procedure for differentiating between normal and abnormal muscle structures. In this study 26 electron micrographs from different myopathies were used with a recognition accuracy of 75%-84% as compared to a human expert.

E. Respiration System

Pneumoconiosis is a disease caused by inhaled dust deposited deep in the lungs. The disease is commonly diagnosed and monitored by chest radiography imaging. Murray *et al.* [67] presented a CAD system for analyzing chest radiographs using multiscale AM-FM representations where a dyadic filterbank, variable spacing, and a local linear phase method were used for estimating the IA and IF magnitude histograms. The proposed framework was evaluated on a rather small database of 22 images giving an area under the receiver operator characteristics curve equal to 1.0.

F. Gynecological

The second most commonly diagnosed cancer, and the most common cancer in women in the world is breast cancer. Mammography is a specialized medical imaging modality for breast imaging using low-energy X-rays. Mammographic breast density is recognized as the strongest predictor to diagnose cancer. Mammograms are classified into one of four density classes based on the Breast Imaging Reporting and Data System (BI-RADS) scale.

On the other hand, a large number of breast cancer cases in women with dense breasts cannot be diagnosed based on mammography. Constantinou *et al.* [68] and [79] applied multiscale AM-FM image analysis using Hilbert based demodulation using Gabor filterbanks. The normalized histogram of the maximum IA across all frequencies for each scale was used to characterize the breast density for each mammogram. The histograms were concatenated to characterize each image and classification into the BI-RADS scale was achieved by comparing the corresponding distribution to the rest using K-NN and Euclidean distance. The proposed methodology was evaluated on 206 mammograms with no abnormalities from the Medical Image Analysis Society mammogram database. The achieved overall correct classifications score was 80%.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

Early detection of breast cancer is related with characterization of architectural distortion in mammograms. In [70], the authors used a combination of fractal analysis and the bidimensional empirical mode decomposition to detect architectural distortions. Based on a dataset of 187 architectural distortion regions extracted from 2191 normal breast parenchyma regions, the method achieved an AUC of 0.95.

In [71], the authors report a classification method to detect breast carcinoma. The proposed approach uses a combination of features that include multilevel iterative variational mode decomposition, Zernike moments, fractal dimension, and entropy features extracted from the VMD components. Based on a dataset of 7,909 breast cancer histological images collected from 82 patients, the authors report classification accuracy rate of 89.61% based on three-fold cross validation.

The second mortality rate in the female population is connected with gynecological cancer. In laparoscopic/hysteroscopic imaging, the physician guides the telescope inside the uterine or abdominal cavity investigating the internal anatomy, in search of suspicious, cancerous lesions. During the exam, the experience of the physician plays a significant role in identifying suspicious ROIs, where in some cases, important ROIs might be ignored and crucial information neglected.

Constantinou *et al.* [72] introduced an adaptive multiscale AM-FM texture analysis using optimal filterbank design to classify normal versus abnormal ROIs of hysteroscopy videos recorded from the endometrium. AM-FM analysis was carried out using Hilbert based demodulation and Gabor filterbanks. The basic concept of this work was the design of a problem - matched filterbank that will facilitate computing the most discriminatory AM-FM feature sets. For each filterbank, the IA, IP and IF estimates were extracted from different scales. Then for each AM-FM component the following texture features were extracted for each ROI: the entropy, median value, and Inter-Quartile Range. The AM-FM feature sets were input into an SVM classifier trained with 50 cases and evaluated with another 50 cases. The highest correct classifications score was 83%.

G. Abdomen

Liver cancer diagnosis using ultrasound imaging is a non-invasive and cheap method but less accurate than surgical biopsy. Acharya *et al.* [73] proposed a computer-aided diagnosis system in order to support specialists in focal liver lesion classification based on the Radon transform and Bidirectional Empirical Mode Decomposition. Statistically significant features were extracted from the DCT coefficients of IMFs. A dataset of 78 normal, 26 benign and 36 malignant images were evaluated where the best accuracy, sensitivity, and specificity obtained were 93%, 91% and 97%, respectively.

Schistosomiasis is a chronic liver diseases leading to hepatic fibrosis, a serious cause of liver failure. Liver histopathology information is essential for fibrosis monitoring and treatment. Ashout *et al.* [74] proposed a new method for microscopy liver tissue image assessment using Empirical Mode Decomposition combined with statistical features. The dataset used was obtained from animal tissue and consisted of 15 cases for each of 3 classes (corresponding to level 1 to 3 of fibrosis) and normal. The best classification accuracy achieved was 98%.

H. Other

In this section, we summarize additional AM-FM applications that are not associated with specific 2D methods applied to different diseases. In [80], the authors investigated the use of AM-FM techniques to detect early stages of malignancies as they appear in optical coherence tomography (OCT). The authors used DCA and the methods described in (12)-(14) to extract the IA, the IP, and the IF to detect scatterers with different size characteristics. The promise of the approach lies in its potential to detect disease early with information extracted from scatterers that remain below the resolution of current OCT systems.

In [81], the authors used a variety of AM-FM techniques to establish the validity of visual and auditory saliency models for processing fMRI videos. For visual silency, the authors used 400 3D Gabor filters for the luminance and color streams to extract motion information in different scales and directions. For the auditory model, they used AM-FM demodulation based on the Teager-Kaiser operator applied over a Gabor filterbank. The approach led to validation of the perceptual relevance of the computational model.

We also list imaging applications that use 1D AM-FM techniques. These applications are distinct from the development of 2D or 3D AM-FM techniques that have been the focus of this paper. In [82], the authors used 1D EMD from resting-fMRI images to study the frequency characteristics of regional homogeneity. In [83], the authors used the 1D CEEMD in fMRI to establish different frequencies of oscillations for patients with depressed Parkinson's disease, non-depressed Parkinson's disease, versus healthy controls. In [84], the authors applied EMD to resting-fMRI images to establish changes in dynamic functional connectivity in early Parkinson's disease. In [85], the authors used CEEMD to automatically decompose the BOLD resting-state fMRI for Mild Cognitive Impairment identification. In [86], the authors used EMD in resting-state fMRI to study their relevance with normal aging In [87], the authors used multivariate Empirical Mode Decomposition to extract task-dependent hemodynamic responses in olfactory-induced fMRI.

VI. Emerging Trends and Possible Future Directions

We begin with a summary of emerging AM-FM methods and applications in section VI-A. We then provide a summary of deep learning methods and discuss connections to the use of AM-FM representations in section VI-B. We provide recommendations for combining deep learning methods with AM-FM methods in section VI-C.

A. Emerging AM-FM Methods and Applications

Interest in medical imaging applications of AM-FM models and methods will surely grow as the methods have matured

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/RBME.2020.2967273, IEEE Reviews in Biomedical Engineering

and several new applications are considered (e.g., see video applications in [88], [89]).

New methods are also being actively developed. As an example, a general purpose segmentation method based on modulation features, generative models, and weighted curve evolution has been described in [32]. The approach consisted of the use of a regularized demodulation algorithm to provide better channel selection for DCA, followed by the use of a generative model based on the DCA output, and the use of a weighted curve evolution to provide excellent segmentation results on the Berkeley Segmentation dataset. Such methods have not been applied to medical imaging yet.

We also have the development of more recent AM-FM methods. For example, iterative filter methods were developed in [90], multi-component AM-FM demodulation in [91], and a generalized mathematical framework for time-frequency analysis in [92]. Additional applications include multi-source image fusion decomposition [93], sparse coding [94], and the use of a 2D Fourier decomposition method for multicomponent AM–FM image analysis [95]. In [96], the authors used an elastic net method to provide a sparse AM-FM image representation that accurately reconstruct the input images.

We also have the development of recent mathematical models that need to be further applied to medical images. First, we expect more applications of the variational mode decomposition methods that have been described in section III-E3 and demonstrated in section IV. Second, we have the Synchrosqueezed Wavelet transform that provides a method to sharpen the time-frequency spread of each EMD harmonic [97]. For images, the synchrosqueezed wave packet transform extended the approach for 2D mode decomposition [98], applied to the curvelet transform in [99], and the monogenic signal in [100]. Yet, we are not aware of any applications of the synchrosqueezed Wavelet transform to medical images.

We also provide a summary of applications of the synchrosqueezed transform. In [101], the authors discussed an application of the synchrosqueezed transform for separating two harmonic functions with different frequencies and amplitudes. An application for separating multi-component AM-FM signals can be found in [102]. An application to determine the painter in different paintings was considered in [103]. An application of 2D synchrosqueezed transforms to analyze atomic crystal images was considered in [104].

B. Deep Learning and Convolutional Neural Networks

The recent emergence of deep learning methods has enabled training and testing complex classifiers on large datasets. In terms of training complex Neural Network systems, the initial breakthrough was reported in recognizing handwritten characters in [105]. Significant progress in the development of deep learning methods is often reported in terms of performance on the ImageNet database of 1,000 categories of 1.2 million (later 1.3M) images [106]. As a brief summary, we have that: (i) AlexNet achieved a top-5 test error of 15.3% with 60M parameters in 2012 [107], followed by (ii) the Inception architecture that reduced the same error to 5.6% [108] in 2016, and (iii) the DenseNet-201 architecture that claimed a 6.34% accuracy with about 20M parameters in 2017 [109].

Deep learning methods have also yielded effective segmentation methods. We have the development of the Fully Convolutional Networks [110], [111] based on transfer learning from ImageNet classifier networks. We then have the SegNet architectures for general image segmentation methods that is based on the use of auto-encoders [112], and the U-net architecture for biomedical image segmentation [113].

For processing medical images and videos, the majority of the deep learning architectures rely heavily on the use convolutional layers, max pooling layers, skip connections, and flattening with fully connected SoftMax layer for the final classification layer. Unlike the use of pre-defined filterbanks for AM-FM analysis methods, deep learning methods need to learn the convolution kernels associated with the different convolutional layers. On the other hand, it is important to note that the advantages of using Gabor filterbanks with neural-networks has been known for a long time (e.g., see [114]). The approach can also be extended to 3D CNN methods using 3D kernels (e.g., 3x3x3). While the individual convolutional layers tend to be very small (e.g., mostly 3x3, 5x5), the large number of interconnections leads to the need to learn several millions of parameters.

Unfortunately, the use of millions of parameters makes CNN hard to interpret and difficult to train. More specifically, each CNN output is input to an activation function (e.g., sigmoid) and then further processed by additional layers. Each convolution layer is actually a biased cross-correlation layer that is maximized when the input is a positive multiple of the cross-correlation kernel. Then, following each convolution layer, the use of max pooling selects the output that is closest to the cross-correlation kernel. It is interesting to note the analogy to the use of the maximum instantaneous amplitude to select the output of each AM-FM channel. In other words, unlike AM-FM methods, convolutional layers effectively select the maximum channel output directly. Beyond the first convolution layer, it is hard to provide meaningful interpretations. For small networks, it is possible to examine the responses of each filter to specific inputs as in the online examples provided in [115] (also see [116]). Undoubtedly, training on large datasets to learn millions of parameters is computationally extremely demanding. Instead, through transfer-learning, new systems can limit training to new applications by re-training existing systems trained on large databases such as ImageNet. As discussed in [117], transfer-learning, even from different modalities, can lead to more robust and improved training than training from scratch on a limited medical database.

Despite the lack of interpretability, CNN-based methods have evolved to address a variety of medical image and video analysis problems. We refer to [118] for a general review of the applications of deep learning methods. We also refer to [1] for a recent survey on the applications of deep learning in medical imaging.

C. Recommendations for Future Work in Deep Learning and AM-FM Methods

We expect to see strong growth in the use of AM-FM representations with deep learning methods. We provide a

summary of our recommendations for combining AM-FM methods with deep learning in terms of classification, feature selection, filterbank design, and transfer learning for medical image analysis applications.

As mentioned earlier, deep learning classifiers are yet to be used with AM-FM features. Clearly, the advantage of using AM-FM features versus raw images comes from the fact that AM-FM features are interpretable. Given the small number of AM-FM features (e.g., histograms of IA and IF) that have already yielded important medical image analysis applications, we expect a significant reduction in the number of parameters that will need to be learned. As an example, [54] showed that FM images can use simple neural network architectures to detect faces, while the same architectures performed extremely poorly on the raw images.

There is a large number of AM-FM features that can benefit from the use of Deep Learning Methods for feature selection. To recognize the problem, we note that an $N \times M$ image decomposed into K components will generate KNM instantaneous amplitude images and $2KNM$ instantaneous frequency components along the x and y directions. Hence, we have $3KNM$ spatially-distributed AM-FM features that can benefit from the combination of classification maps and Convolutional Neural Nets as described in [119]. To explain the approach, we assume the use of global average pooling (GAP) over the $3KNM$ AM-FM feature maps (averages over each map) that are input to a fully-connected softmax layer. The classification map is defined as the weighted average of the feature maps where the weights come from the weights of the GAP layer associated with the classification category. Thus, the classification maps allow us to visualize the AM-FM features that are used for the classification. In turn, this approach can support future research in interpretable AM-FM feature selection where the dominating features can be used for representing the underlying textures.

Interpretable filterbank design can also benefit from the use of classification maps. As an example, a clustering of the most significant instantaneous frequency vectors, extracted from the class action maps, can be used to design filters that can provide strong coverage of the 2D frequency spectrum that contain them.

Assuming pre-training of AM-FM features with deep models over large databases, future medical imaging applications can benefit from transfer learning. Just like Convolutional Neural Networks, AM-FM filterbanks can be applied to any input image size. Then, future medical imaging applications will simply need to replace and retrain the top layers for specific applications.

VII. Conclusion

Over the years, substantial progress has been made on the development of AM-FM demodulation methods. Several of the most dominant methods have been summarized in section III. While AM-FM demodulation methods are still of interest, we now have robust methods that have proven effective in many applications. Yet, there are clearly many important new methods that have not been applied to medical image or medical video analysis.

In general, there is a need to consider medical image analysis using a multitude of different AM-FM features. To recognize this need, note that AM-FM models and transforms can generate a large number of AM-FM components. Yet, in terms of applications, the majority of the applications are concerned with the use of a single AM-FM component. More recently, there has been great interest in the use of multiscale AM-FM models that extract components from each scale. Using multiscale representations, the approach generates AM-FM components and features from each scale. Yet, even when using multiscale methods, the majority of the applications only consider histograms of the IF, the IA, the IF magnitude. The full 2D AM-FM components have not been seriously considered.

Future research could consider the use of convolutional neural networks with AM-FM components and features. In such applications, the goal would be to set the input channels to the individual AM-FM components that will be input to a CNN with a significantly reduced number of parameters. In other words, the AM-FM components and features will act as the low-level features that are typically extracted through a large-scale training efforts. The use of AM-FM components and features will provide much better interpretability with significantly less-complex neural network architectures.

References

[1] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[2] B. Goodman and S. Flaxman, "European union regulations on algorithmic decision-making and a "right to explanation"," *AI Magazine*, vol. 38, no. 3, pp. 50–57, 2017.

[3] V. Andrearczyk and P. F. Whelan, "Using filter banks in convolutional neural networks for texture classification," *Pattern Recognition Letters*, vol. 84, pp. 63–69, 2016.

[4] N. Liu, M. Zhang, H. Li, Z. Sun, and T. Tan, "Deepiris: Learning pairwise filter bank for heterogeneous iris verification," *Pattern Recognition Letters*, vol. 82, pp. 154–161, 2016.

[5] M. Clark, A. C. Bovik, and W. S. Geisler, "Texture segmentation using Gabor modulation/demodulation," *Pattern Recognit. Lett.*, vol. 6, no. 4, pp. 261–267, 1987.

[6] M. Pattichis, C. Pattichis, M. Avraam, A. Bovik, and K. Kyriacou, "AM-FM texture segmentation in electron microscopic muscle imaging," *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1253–1257, 2000.

[7] C. Agurto, V. Murray, E. Barriga, S. Murillo, M. Pattichis, H. Davis, S. Russell, M. Abramoff, and P. Soliz, "Multiscale am-fm methods for diabetic retinopathy lesion detection," *IEEE Transactions on Medical Imaging*, vol. 29, no. 2, pp. 502–512, Feb 2010.

[8] M. Pattichis and A. Bovik, "Analyzing image structure by multidimensional frequency modulation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 5, pp. 753–766, May 2007.

[9] ——, "Analyzing image structure by multidimensional frequency modulation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 5, pp. 753–766, May 2007.

[10] V. Murray, M. S. Pattichis, E. S. Barriga, and P. Soliz, "Recent multiscale AM-FM methods in emerging applications in medical imaging," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, Feb 2012.

[11] M. S. Pattichis, G. Panayi, A. C. Bovik, and S.-P. Hsu, "Fingerprint classification using an AM-FM model," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 951–954, 2001.

[12] M. S. Pattichis and A. C. Bovik, "AM-FM expansions for images," in *8th European Signal Processing Conference. EUSIPCO 1996.* IEEE, 1996, pp. 1–4.

[13] M. S. Pattichis, "AM-FM transforms with applications," Ph.D. dissertation, The University of Texas at Austin, 5 1998.

[14] M. S. Pattichis, A. C. Bovik, J. W. Havlicek, and N. D. Sidiropoulos, "Multidimensional orthogonal FM transforms," *IEEE Transactions on Image Processing*, vol. 10, no. 3, pp. 448–464, 2001.

[15] N. D. Sidiropoulos, M. S. Pattichis, A. C. Bovik, and J. W. Havlicek, "Coperm: Transform-domain energy compaction by optimal permutation," *IEEE transactions on signal processing*, vol. 47, no. 6, pp. 1679–1688, 1999.

[16] V. Murray, P. Rodriguez, and M. S. Pattichis, "Multiscale AM-FM demodulation and image reconstruction methods with improved accuracy," *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1138–1152, May 2010.

[17] J. P. Havlicek, "AM-FM image models," Ph.D. dissertation, University of Texas at Austin, 1996.

[18] T. Bulow and G. Sommer, "Hypercomplex signals-a novel extension of the analytic signal to the multidimensional case," *IEEE Transactions on signal processing*, vol. 49, no. 11, pp. 2844–2852, 2001.

[19] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, Dec 2001.

[20] ——, "The monogenic scale-space: A unifying approach to phase-based image processing in scale-space," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 1, pp. 5–26, Jul 2004.

[21] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "Multidimensional quasi-eigenfunction approximations and multicomponent am-fm models," *IEEE Transactions on Image Processing*, vol. 9, no. 2, pp. 227–242, 2000.

[22] J. P. Havlicek, P. C. Tay, and A. C. Bovik, "AM-FM image models: Fundamental techniques and emerging trends," in *Handbook of Image and Video Processing*. Elsevier, 2005, pp. 377–395.

[23] A. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by gabor wavelets," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 691–712, Mar 1992.

[24] A. Bovik, P. Maragos, and T. Quatieri, "AM-FM energy detection and separation in noise using multiband energy operators," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3245–3265, 1993.

[25] J. P. Havlicek, D. S. Harding, and A. C. Bovik, "The multicomponent am-fm image representation," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 1094–1100, 1996.

[26] J. Havlicek and A. Bovik, "Image modulation models," *Handbook of image and video processing*, pp. 305–316, 2000.

[27] C. T. Nguyen and J. P. Havlicek, "Am-fm image filters," in *IEEE International Conference on Image Processing*. IEEE, 2008, pp. 789–792.

[28] E. Simoncelli, W. Freeman, E. Adelson, and D. Heeger, "Shiftable multiscale transforms," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 587–607, Mar 1992.

[29] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *International Conference on Image Processing*, vol. 3. IEEE, 1995, pp. 444–447.

[30] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis," *IEEE transactions on signal processing*, vol. 41, no. 10, pp. 3024–3051, 1993.

[31] P. Maragos and A. C. Bovik, "Image demodulation using multidimensional energy separation," *Journal of the Optical Society of America A*, vol. 12, no. 9, p. 1867, Sep 1995.

[32] I. Kokkinos, G. Evangelopoulos, and P. Maragos, "Texture analysis and segmentation using modulation features, generative models, and weighted curve evolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 142–157, Jan 2009.

[33] F. Salzenstein, A. O. Boudraa, and J. C. Cexus, "Generalized higher-order nonlinear energy operators," *J. Opt. Soc. Am. A*, vol. 24, no. 12, pp. 3717–3727, Dec 2007.

[34] E. H. S. Diop, A. O. Boudraa, and F. Salzenstein, "Higher order teager-kaiser operators for image analysis: Part i - a monocomponent image demodulation," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, Apr 2009.

[35] E. H. S. Diop and A. O. Boudraa, "Higher order teager-kaiser operators for image analysis: PART II - a multicomponent image demodulation," in *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE, Nov 2009.

[36] A. O. Boudraa and F. Salzenstein, "Teager–kaiser energy methods for signal and image analysis: A review," *Digital Signal Processing*, vol. 78, pp. 338–375, Jul 2018.

[37] W. Liu and B. Santhanam, "Wideband image demodulation via bi-dimensional multirate frequency transformations," *Journal of the Optical Society of America A*, vol. 33, no. 9, p. 1668, Aug 2016.

[38] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, Mar 1998.

[39] N. E. Huang, M.-L. C. Wu, S. R. Long, S. S. P. Shen, W. Qu, P. Gloersen, and K. L. Fan, "A Confidence Limit for the Empirical Mode Decomposition and Hilbert Spectral Analysis," *Proc. R. Soc. London{,} Ser. A*, vol. 459, pp. 2317–2345, 2003.

[40] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 112–114, Feb 2004.

[41] R. J. Gledhill, "Methods for investigating conformational change in biomolecular simulations," Ph.D. dissertation, University of Southampton, 2003.

[42] Z. Wu and N. E. Huang, "Ensemble Empirical Mode Decomposition: A noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 01, no. 01, pp. 1–41, Jan 2009.

[43] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, May 2011.

[44] N. E. Huang, "Computer implemented empirical mode decomposition method, apparatus and article of manufacture," Nov 1999.

[45] Z. Wu, N. E. Huang, and X. Chen, "The multi-dimensional ensemble empirical mode decomposition method," *Advances in Adaptive Data Analysis*, vol. 01, no. 03, pp. 339–372, Jul 2009.

[46] A. H. Heurtier, G. Mahe, and P. Abraham, "Multi-dimensional complete ensemble empirical mode decomposition with adaptive noise applied to laser speckle contrast images," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 2103–2117, Oct 2015.

[47] J. Nunes, Y. Bouaoune, E. Deléchelle, O. Niang, and P. Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image and Vision Computing*, vol. 21, no. 12, pp. 1019–1026, Nov 2003.

[48] J. C. Nunes, S. Guyot, and E. Deléchelle, "Texture analysis based on local analysis of the bidimensional empirical mode decomposition," *Machine Vision and applications*, vol. 16, no. 3, pp. 177–188, 2005.

[49] K. Dragomiretskiy and D. Zosso, "Two-dimensional variational mode decomposition," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, X. C. Tai, E. Bae, T. F. Chan, and M. Lysaker, Eds. Cham: Springer International Publishing, 2015, pp. 197–208.

[50] ——, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, Feb 2014.

[51] D. Zosso, K. Dragomiretskiy, A. L. Bertozzi, and P. S. Weiss, "Two-dimensional compact variational mode decomposition," *Journal of Mathematical Imaging and Vision*, vol. 58, no. 2, pp. 294–320, 2017.

[52] J. C. Nunes and E. Deléchelle, "Empirical Mode Decomposition: Applications on Signal and Image Processing," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 125–175, Jan 2009.

[53] D. Boutte and B. Santhanam, "A hybrid ica-svm approach to continuous phase modulation recognition," *IEEE Signal Processing Letters*, vol. 16, no. 5, pp. 402–405, 2009.

[54] L. S. Tapia, "The importance of the instantaneous phase in detecting faces with convolutional neural networks," Master's thesis, Department of Electrical and Computer Engineering, The University of New Mexico, 2019.

[55] C. P. Loizou, V. Murray, M. S. Pattichis, M. Pantziaris, and C. S. Pattichis, "Multiscale amplitude-modulation frequency-modulation (AM–FM) texture analysis of ultrasound images of the intima and media layers of the carotid artery," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 2, pp. 178–188, Mar 2011.

[56] F. Molinari, U. Raghavendra, A. Gudigar, K. M. Meiburger, and U. Rajendra Acharya, "An efficient data mining framework for the characterization of symptomatic and asymptomatic carotid plaque using bidimensional empirical mode decomposition technique," *Med. Biol. Eng. Comput.*, vol. 56, no. 9, pp. 1579–1593, 2018.

[57] T. T. Tran, V. T. Pham, C. Lin, H. W. Yang, Y. H. Wang, K. K. Shyu, W. Y. I. Tseng, M. Y. M. Su, L. Y. Lin, and M. T. Lo, "Empirical Mode Decomposition and Monogenic Signal based Approach for Quantification of Myocardial Infarction from MR Images," *IEEE J. Biomed. Heal. Informatics*, vol. PP, no. c, p. 1, 2018.

[58] C. P. Loizou, V. Murray, M. M. Pattichis, I. Seimenis, M. Pantziaris, and C. S. Pattichis, "Multiscale amplitude-modulation frequency-modulation (AM–FM) texture analysis of multiple sclerosis in brain MRI images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 1, pp. 119–129, Jan 2011.

[59] S. P. Washimkar and S. D. Chede, "Prediction of multiple sclerosis in brain mri images using hybrid segmentation," in *2017 International Conference on Signal Processing and Communication (ICSPC)*, July 2017, pp. 234–239.

[60] A. Neubauer, A. M. Tomé, A. Kodewitz, J. Górriz, C. G. Puntonet, and E. W. Lang, "Bidimensional ensemble empirical mode decomposition of functional biomedical images," *Advances in Adaptive Data Analysis*, vol. 6, no. 01, p. 1450004, Jan 2014.

[61] A. Rojas, J. M. Górriz, J. Ramírez, I. A. Illán, F. J. Martínez-Murcia, A. Ortiz, M. Gómez Río, and M. Moreno-Caballero, "Application of Empirical Mode Decomposition (EMD) on DaTSCAN SPECT images to explore Parkinson Disease," *Expert Syst. Appl.*, vol. 40, no. 7, pp. 2756–2766, 2013.

[62] C. Agurto, E. S. Barriga, V. Murray, S. Nemeth, R. Crammer, W. Bauman, G. Zamora, M. S. Pattichis, and P. Soliz, "Automatic detection of diabetic retinopathy and age-related macular degeneration in digital fundus images," *Investigative Ophthalmology & Visual Science*, vol. 52, no. 8, p. 5862, 2011.

[63] C. Agurto, V. Murray, H. Yu, J. Wigdahl, M. Pattichis, S. Nemeth, E. S. Barriga, and P. Soliz, "A multiscale optimization approach to detect exudates in the macula," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1328–1336, Jul 2014.

[64] C. Agurto, H. Yu, V. Murray, M. S. Pattichis, S. Nemeth, S. Barriga, and P. Soliz, "A multiscale decomposition approach to detect abnormal vasculature in the optic disc," *Computerized Medical Imaging and Graphics*, vol. 43, pp. 137–149, Jul 2015.

[65] S. Lahmiri and A. Shmuel, "Variational mode decomposition based approach for accurate classification of color fundus images with hemorrhages," *Opt. Laser Technol.*, vol. 96, pp. 243–248, 2017.

[66] S. Maheshwari, R. B. Pachori, V. Kanhangad, S. V. Bhandary, and U. R. Acharya, "Iterative variational mode decomposition based automated detection of glaucoma using fundus images," *Computers in biology and medicine*, vol. 88, pp. 142–149, 2017.

[67] V. Murray, M. Pattichis, and P. Soliz, "Retrieval and classification of pneumoconiosis chest radiograph images using multiscale am-fm methods," in *Forty-Third Asilomar Conference on Signals, Systems and Computers*. IEEE, 2009, pp. 12–16.

[68] I. Constantinou, M. Pattichis, C. Tziakouri, C. Pattichis, S. Petroudi, and C. Nicosia, "Multiscale am-fm models and instantaneous amplitude evaluation for mammographic density classification," *MIUA*, pp. 271–276, 2014.

[69] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok *et al.*, "The mammographic image analysis society digital mammogram database," in *Exerpta Medica. International Congress Series*, vol. 1069, 1994, pp. 375–378.

[70] I. Zyout and R. Togneri, "A computer-aided detection of the architectural distortion in digital mammograms using the fractal dimension measurements of BEMD," *Comput. Med. Imaging Graph.*, vol. 70, no. April, pp. 173–184, 2018.

[71] S. Chattoraj and K. Vishwakarma, "Classification of histopathological breast cancer images using iterative vmd aided zernike moments & textural signatures," *arXiv preprint arXiv:1801.04880*, 2018.

[72] I. Constantinou, M. Pattichis, V. Tanos, M. Neofytou, and C. Pattichis, "An adaptive multiscale AM-FM texture analysis system with application to hysteroscopy imaging," in *2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE)*. IEEE, Nov 2012.

[73] U. R. Acharya, J. E. W. Koh, Y. Hagiwara, J. H. Tan, A. Gertych, A. Vijayananthan, N. A. Yaakup, B. J. J. Abdullah, M. K. B. M. Fabell, and C. H. Yeong, "Automated diagnosis of focal liver lesions using bidirectional empirical mode decomposition features," *Computers in biology and medicine*, vol. 94, pp. 11–18, 2018.

[74] D. S. Ashour, D. M. Abou Rayia, M. Maher Ata, A. S. Ashour, and M. M. Abd Elnaby, "Hybrid feature extraction techniques for microscopic hepatic fibrosis classification," *Microsc. Res. Tech.*, vol. 81, no. 3, pp. 338–347, 2018.

[75] M. J. Leahy, *Microcirculation imaging*. John Wiley & Sons, 2012.

[76] A. Humeau-Heurtier, P. Abraham, and G. Mahe, "Analysis of laser speckle contrast images variability using a novel empirical mode decomposition: Comparison of results with laser doppler flowmetry signals variability," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 618–627, Feb 2015.

[77] V. Murray, M. Pattichis, and P. Soliz, "New am-fm analysis methods for retinal image characterization," in *42nd Asilomar Conference on Signals, Systems and Computers*. IEEE, 2008, pp. 664–668.

[78] E. Barriga, V. Murray, C. Agurto, M. Pattichis, S. Russell, M. Abramoff, H. Davis, and P. Soliz, "Multi-scale AM-FM for lesion phenotyping on age-related macular degeneration," in *2009 22nd IEEE International Symposium on Computer-Based Medical Systems*. IEEE, Aug 2009.

[79] S. Petroudi, I. Constantinou, M. Pattichis, C. Tziakouri, K. Marias, and C. Pattichis, "Evaluation of spatial dependence matrices on multiscale instantaneous amplitude for mammogram classification," in *IFMBE Proceedings*. Springer International Publishing, 2015, pp. 156–159.

[80] C. Pitris, A. Kartakoullis, and E. Bousi, "AM-FM techniques in the analysis of optical coherence tomography signals," *Journal of Biophotonics*, vol. 2, no. 6-7, pp. 364–369, Jul 2009.

[81] G. Panagiotaropoulou, P. Koutras, A. Katsamanis, P. Maragos, A. Zlatintsi, A. Protopapas, E. Karavasilis, and N. Smyrnis, "Fmri-based perceptual validation of a computational model for visual and auditory saliency in videos," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, pp. 699–703.

[82] X. Song, Y. Zhang, and Y. Liu, "Frequency specificity of regional homogeneity in the resting-state human brain," *PLoS One*, vol. 9, no. 1, 2014.

[83] L. Qian, Y. Zhang, L. Zheng, X. Fu, W. Liu, Y. Shang, Y. Zhang, Y. Xu, Y. Liu, H. Zhu *et al.*, "Frequency specific brain networks in parkinson's disease and comorbid depression," *Brain imaging and behavior*, vol. 11, no. 1, pp. 224–239, 2017.

[84] D. Cordes, X. Zhuang, M. Kaleem, K. Sreenivasan, Z. Yang, V. Mishra, S. J. Banks, B. Bluett, and J. L. Cummings, "Advances in functional magnetic resonance imaging data analysis methods using Empirical Mode Decomposition to investigate temporal changes in early Parkinson's disease," *Alzheimer's Dement. Transl. Res. Clin. Interv.*, vol. 4, pp. 372–386, 2018.

[85] L. Qian, L. Zheng, Y. Shang, Y. Zhang, and Y. Zhang, "Intrinsic frequency specific brain networks for identification of MCI individuals using resting-state fMRI," *Neuroscience Letters*, vol. 664, pp. 7 – 14, 2018.

[86] A. C. Yang, S.-J. Tsai, C.-P. Lin, C.-K. Peng, and N. E. Huang, "Frequency and amplitude modulation of resting-state fmri signals and their functional relevance in normal aging," *Neurobiology of aging*, vol. 70, pp. 59–69, 2018.

[87] K. W. Wang, C. Y. Chen, H. H. Chang, C. C. Hsu, G. Y. Lan, H. T. Hsu, K. K. Shyu, W. P. Chan, and P. L. Lee, "A Multivariate Empirical Mode Decomposition–Based Data-Driven Approach for Extracting Task-Dependent Hemodynamic Responses in Olfactory-Induced fMRI," *IEEE Access*, vol. 7, pp. 15 375–15 388, 2019.

[88] W. Shi, M. S. Pattichis, S. Celedón-Pattichis, and C. LópezLeiva, "Robust head detection in collaborative learning environments using am-fm representations," in *2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, April 2018, pp. 1–4.

[89] W. Shi, M. Pattichis, S. Celedon-Pattichis, and C. LopezLeiva, "Dynamic group interactions in collaborative learning videos," in *2018 Asilomar Conference on Signals, Systems and Computers*. IEEE, Nov 2018, pp. 1528–1531.

[90] A. Cicone and H. Zhou, "Multidimensional iterative filtering method for the decomposition of high–dimensional non–stationary signals," *Numerical Mathematics: Theory, Methods and Applications*, vol. 10, no. 2, pp. 278–298, 2017.

[91] A. K. George and P. Sumathi, "Pre-filtered phase-locking scheme for multi-component am–fm signal decomposition," *Circuits, Systems, and Signal Processing*, vol. 37, no. 2, pp. 752–769, 2018.

[92] S. Sandoval and P. L. De Leon, "The instantaneous spectrum: A general framework for time-frequency analysis," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5679–5693, 2018.

[93] W. Huijuan, J. Yong, and M. Xingmin, "Fast bi-dimensional empirical mode based multisource image fusion decomposition," in *2019 28th Wireless and Optical Communications Conference (WOCC)*, May 2019, pp. 1–4.

[94] E. H. S. Diop and K. Skretting, "Frequency separation method based on sparse coding," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 5192–5196.

[95] P. Singh and S. D. Joshi, "Some Studies on Multidimensional Fourier Theory for Hilbert Transform, Analytic Signal and AM–FM Representation," *Circuits, Systems, and Signal Processing*, pp. 1–28, 2019.

[96] I. Constantinou, M. S. Pattichis, and C. S. Pattichis, "Multiscale AM-FM image reconstructions based on elastic net regression and gabor filterbanks," in *2013 Asilomar Conference on Signals, Systems and Computers*. IEEE, Nov 2013, pp. 1985–1989.

[97] I. Daubechies, J. Lu, and H. T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and*

Computational Harmonic Analysis, vol. 30, no. 2, pp. 243–261, Mar 2011.

[98] H. Yang and L. Ying, "Synchrosqueezed wave packet transform for 2d mode decomposition," *SIAM Journal on Imaging Sciences*, vol. 6, no. 4, pp. 1979–2009, 2013.

[99] ——, "Synchrosqueezed curvelet transform for 2d mode decomposition," *arXiv preprint arXiv:1310.6079*, 2013.

[100] M. Clausel, T. Oberlin, and V. Perrier, "The monogenic synchrosqueezed wavelet transform: a tool for the decomposition/demodulation of am–fm images," *Applied and Computational Harmonic Analysis*, vol. 39, no. 3, pp. 450–486, 2015.

[101] H. T. Wu, P. Flandrin, and I. Daubechies, "One or two frequencies? The synchrosqueezing answers," *Advances in Adaptive Data Analysis*, vol. 03, no. 01n02, pp. 29–39, Apr 2011.

[102] S. Meignen, T. Oberlin, and S. McLaughlin, "A new algorithm for multicomponent signals analysis based on SynchroSqueezing: With an application to signal sampling and denoising," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5787–5798, Nov 2012.

[103] H. Yang, J. Lu, W. P. Brown, I. Daubechies, and L. Ying, "Quantitative canvas weave analysis using 2-d synchrosqueezed transforms: Application of time-frequency analysis to art investigation," *Signal Processing Magazine, IEEE*, 2015.

[104] H. Yang, J. Lu, and L. Ying, "Crystal image analysis using 2d synchrosqueezed transforms," *Multiscale Modeling & Simulation*, vol. 13, no. 4, pp. 1542–1572, 2015.

[105] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[106] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[107] I. Sutskever, G. E. Hinton, and A. Krizhevsky, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[108] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[109] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[110] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[111] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *arXiv preprint arXiv:1605.06211*, 2016.

[112] V. Badrinarayanan, A. Handa, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling," *arXiv preprint arXiv:1505.07293*, 2015.

[113] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[114] J. G. Daugman, "Complete discrete 2-d gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 36, no. 7, pp. 1169–1179, 1988.

[115] Y. LeCun, "Lenet-5, convolutional neural networks," http://yann.lecun.com/exdb/lenet/, accessed: 2019-06-13.

[116] A. E. U. Cerna, M. Pattichis, D. P. vanMaanen, L. Jing, A. A. Patel, J. V. Stough, C. M. Haggerty, and B. K. Fornwalt, "Interpretable neural networks for predicting mortality risk using multi-modal electronic health records," *arXiv preprint arXiv:1901.08125*, 2019.

[117] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1299–1312, 2016.

[118] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[119] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2921–2929.

Kyriacos P. Constantinou received his B.Sc. from the Hellenic Army Academy, Athens, Greece, in 2002 and the M.Sc. degree in Computer Science from the University of Cyprus, Nicosia, Cyprus in 2012. He is currently working towards the Ph.D. degree at the Department of Computer Science, University of Cyprus.

His research interests include medical image and video analysis, and computer vision.

Ioannis P. Constantinou received his Dipl-Ing. Degree in Electrical Engineering and Computer Engineering from National Technical University of Athens, Greece, his MSc in Biomedical Engineering from the Postgraduate Program of the University of Patras and National Technical University of Athens, and his PhD degree from the Department Computer Science of the University of Cyprus in the area of image analysis/computer vision.

Dr. Constantinou has been involved in several EU-funded and RPF-funded research projects related to robotics, telemedicine, e-learning, Internet of Things, advance image processing and analysis, biological procedure modelling and simulation and e-health. He contributed to more than 20 papers in scientific conference and journals in the area of robotics, medical image analysis, computer vision and ehealth.

Constantinos S. Pattichis (S'88-M'88-SM'99-F'18) is currently Professor with the Department of Computer Science of the University of Cyprus. His research interests include ehealth and mhealth, medical imaging, biosignal analysis, life sciences informatics, and intelligent systems. He has been involved in numerous projects in these areas funded by EU, the National Research Foundation of Cyprus, the INTERREG and other bodies. He is a Fellow of IEEE, Fellow of IET, Fellow of the European Alliance of Medical and Biolological Sciences (EAAMBES) and Fellow of the International Academy of Medical and Biological Engineering (IAMBE).

Marios S. Pattichis (M'99–SM'06) received the B.Sc. degree (High Hons. and Special Hons.) in computer sciences, the B.A. degree (High Hons.) in mathematics, the M.S. degree in electrical engineering, and the Ph.D. degree in computer engineering from The University of Texas at Austin, Austin, TX, USA, in 1991, 1991, 1993, and 1998, respectively.

He is currently a Gardner Zemke Professor and Associate Chair with the Department of Electrical and Computer Engineering and a fellow with the Center for Collaborative Research and Community Engagement in the College of Education, at The University of New Mexico (UNM), Albuquerque, NM, USA. At UNM, he is currently the Director of the Image and Video Processing and Communications Lab (ivPCL). His current research interests include digital image and video processing, educational models for learning, video communications, dynamically reconfigurable hardware architectures, and biomedical, educational and space image-processing applications. For his development of the Digital Logic Design Labs at UNM, he was recognized by the Xilinx Corporation in 2003 and the UNM School of Engineering's Harrison Faculty Excellent Award in 2006. He is the general co-chair of the 2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), and served as the general chair for the 2008 SSIAI. He was a founding Co-PI of the COSMIAC Research Center, UNM. He is currently a Senior Associate Editor of the IEEE Transactions on Image Processing. He has served as a Senior Associate Editor for the IEEE Signal Processing Letters, an Associate Editor for the IEEE Transactions on Image Processing and the IEEE Transactions on Industrial Informatics, and a Guest Associate Editor for the IEEE Transactions on Information Technology in Biomedicine.