Physically Plausible Adaptive Variable Density Sampling Scheme based on k-ABC

Akshay Kumar Jagadish[a], Kasi Rajgopal[b]

[a] *Electrical and Electronics Engineering, NITK, Surathkal.*
[b] *Electrical Engineering, Indian Institute Of Science, Bangalore.*

Abstract

Swarm optimization algorithms which are motivated by the hierarchical working, efficient self-organizing skills and highly developed foraging ability of the bee population are being increasingly used in wide array fields. In this paper, we propose a novel data-driven algorithm, using swarm intelligence of artificial bee colony (ABC), called k-ABC, to generate an adaptive variable density sampling (VDS) scheme for compressive sampling (CS) based data acquisition for fast MR imaging. The algorithm exploits the behaviour of the three types of bees - scout-bees, employed-bees and onlooker bees, with certain modifications based on the characteristics of variable density sampling. We introduce the concept of searching for the high quality food sources in annular regions of varying widths, called as bins, to optimise the process of foraging. The k-ABC algorithm uses magnitude k-space distribution of a reference single-slice MR image as the underlying fitness value distribution to generate an adaptive sampling scheme. We have also addressed the problem of designing a tailor-made template sampling scheme for 3D MR volume imaging for the very first time. Retrospective simulations show that the proposed k-ABC adaptive VDS scheme gives significant improvement over other sampling schemes for both single slice and multi-slice MR imaging. Further, for the task of implementation, a modified projection algorithm that takes into account the location of each sample in k-space has been introduced, which provides a significant improvement in the reconstructed image quality with minimum trade-off in terms of scan-time.

Keywords: Artificial Bee Colony (ABC), k-ABC, Compressive sampling (CS),

Artificial Bee Colony (ABC) based Variable Density Sampling Scheme for CS-MRI

Akshay Kumar Jagadish[†], Soumya Goswami[*], Pramit Saha[*], Satrajit Chakrabarty[*] and Kasi Rajgopal[‡]
[†] Electrical and Electronics Engineering, NITK, Surathkal. [*]Electrical Engineering, Jadavpur University, Kolkata.
[‡]Electrical Engineering, Indian Institute Of Science, Bangalore.

***Abstract*—The self-sustained dynamics of the bee population in nature is a result of their hierarchical working culture, efficient organizing skills and unique highly developed foraging ability, which enables them to interact effectively among each other as well as with their environment. In this paper, a novel algorithm utilizing the bee's swarm intelligence, and its heuristics based on quality and quantity of food sources (nectars) is proposed to generate a variable density sampling (VDS) scheme for compressive sampling (CS) based fast MRI data acquisition. The algorithm uses the scout-bees for global random selection process which is further fine-tuned by employed and onlooker-bees who forage locally in the neighborhood giving prime importance to points possessing high fitness values (or high energy) usually located around the center of *k*-space. The algorithm introduces the concept of searching for the high quality food sources in annular regions, called as bins, of varying widths. Retrospective CS-MRI simulations show that the proposed *k*-ABC based VDS scheme performs significantly better than other sampling schemes.**

***Index Terms*—Artificial Bee Colony (ABC), k-ABC, Compressive sampling (CS), Variable Density Sampling (VDS).**

I. Introduction

Compressed Sensing (CS) [1], a sparse signal recovery technique, is being widely used to reconstruct MR images because unlike traditional Nyquist sampling method, it requires only few measurements provided the underlying image has sparse representation in some transform domain (like Fourier domain), and the acquired samples are incoherent in the same domain. Thereby, showing tremendous potential in reducing the data acquisition time.

With the application of compressed sensing to rapid and high quality MRI, Lustig et al. [2] developed the idea of randomized Cartesian sub-sampling. The random sampling concept was extended to draw samples from a polynomial based probability density function (pdf) using Monte-Carlo algorithm with minimum peak interference. However, constructing a pdf that will generate a suitable sampling scheme with appropriate number of low and high frequency k-space samples is a challenging problem. Ravishankar et al. [3] have proposed an adaptive sampling algorithm using a novel sparsifying Dictionary Learning technique which simultaneously learns an image patch based dictionary and reconstructs the image using under-sampled data iteratively.

Traveling Salesman Problem (TSP) based optimization for sampling scheme generation has been proposed by Chauffert et al. [4]. In this method, a continuous k-space trajectory is traced using sample points drawn from a specific target distribution (π-distribution) by solving TSP. Recently, Adcock et. al [5] have proposed the usage of optimal sub-sampling strategy based on the structure of the signal, instead of uniform random sub-sampling which yields poor results.

Artificial bee colony (ABC) algorithm is a swarm intelligence algorithm, motivated by the foraging behavior of bee swarms. Karaboga et al. [6] developed an idea based on honey bee swarm for numerical optimization and later proposed a combinatorial ABC algorithm for solving traveling salesman problem.

In this paper, we have designed a meta-heuristic iterative artificial bee colony algorithm to mimic a VDS scheme for fast MR imaging. It is shown that the proposed novel k-ABC based sampling scheme gives significant improvement over other well-known algorithms for drawing samples from the k-space.

II. k-ABC Algorithm

Motivation:

In nature, the nectar collected by the bees are stored in cells of the hive. Similarly, in MR imaging, sampling points or samples that are closer to center of k-space hold the essential features pertaining to overall structure of the image, contrast details and signal to noise ratio. Hence, certain region in the center of k-space, possessing high energy [7], is sampled densely which is analogous to the bee-hive. The concept of variation in the quantity and quality of food sources gradually with distance from the hive has been used to design the k-space sampling scheme. In order to preserve randomness in sampling scheme necessary for compressed sensing based reconstruction, random vectors or agents called scout-bees have been utilized for foraging for best nectars globally. Employed and onlooker-bees are made to forage locally in the neighborhood so as to further fine-tune the components possessing higher fitness values that are usually concentrated close to the center of k-space. Thus, a meta-heuristic algorithm has been developed given the better performance of heuristics [4] as follows:

A. Random Search by the Scout Bees

The hive of the bee system is the center or source of bee-movement. In this work, the beehive is initialized at the center of k-space representing the high energy low frequency peak of the k-space. A specified number of scout-bees perform random search and convey the information of location of food sources to unemployed-bees.

Selfie Detection by Synergy-Constriant Based Convolutional Neural Network

Yashas Annadani, Vijaykrishna Naganoor, Akshay Kumar Jagadish and Krishnan Chemmangat
Electrical and Electronics Engineering, NITK-Surathkal, India.

***Abstract*—Categorisation of huge amount of data on the multimedia platform is a crucial task. In this work, we propose a novel approach to address the subtle problem of selfie detection for image database segregation on the web, given rapid rise in the number of selfies being clicked. A Convolutional Neural Network (CNN) is modeled to learn a synergy feature in the common subspace of head and shoulder orientation, derived from Local Binary Pattern (LBP) and Histogram of Oriented Gradients (HOG) features respectively. This synergy was captured by projecting the aforementioned features using Canonical Correlation Analysis (CCA). We show that the resulting networks convolutional activations in the neighbourhood of spatial keypoints captured by SIFT are discriminative for selfie-detection. In general, proposed approach aids in capturing intricacies present in the image data and has the potential for usage in other subtle image analysis scenarios apart from just selfie detection. We investigate and analyse the performance of the popular CNN architectures (GoogleNet, Alexnet), used for other image classification tasks, when subjected to the task of detecting the selfies on the multimedia platform. The results of the proposed approach are compared with these popular architectures on a dataset of ninety thousand images comprising of roughly equal number of selfies and non-selfies. Experimental results on this dataset shows the effectiveness of the proposed approach.**

***Index Terms*—Selfie; Deep Learning; Convolutional Neural Networks; Canonical Correlation Analysis**

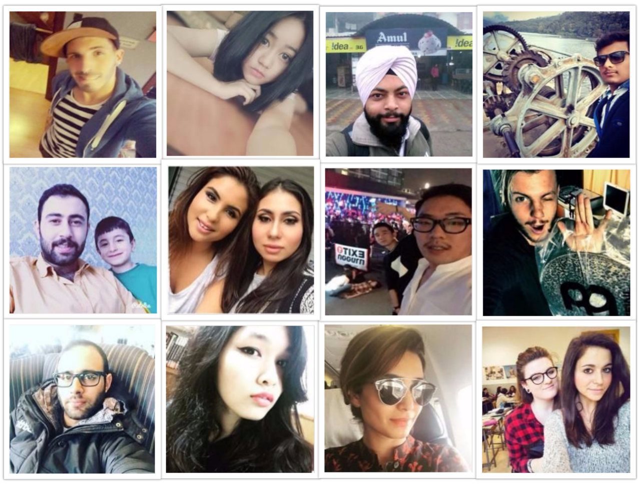

Fig. 1. Sample images from the selfie dataset [10].

I. Introduction

Self-Portraits, popularly known as Selfies have become ubiquitous over the past five years, with the burgeoning of social network and photo sharing platforms. Manually annotating the images as self portraits would be impossible given the massive volume of images entering the web. On the other hand, the tags that come along with these images are either not available always or not reliable, if present. In this milieu, it becomes extremely important to design an efficient method that classifies a given image as selfie or not as it has profound applications in sentiment analysis [1] [2], large scale image database segregation and retrieval [3], and other psychological studies [4] [5]. Apart from the aforementioned applications, it can also be used in scene understanding as people take selfies not just as a picture of oneself, but to capture the essence of the background, like a historic monument, that might exist in the background.

We have hence formulated selfie detection as a binary classification problem where an image has to classified as a selfie or not a selfie and address this problem using Convolutional Neural Networks (CNN), which capture view invariant feature representations and have become successful in the recent past on many image classification tasks like object recognition [6] [7], object detection [8] and face recognition [9].

Selfie detection is not a straightforward problem. To address this problem, we need to answer the following impending questions: What characterizes an image as a selfie? How do computer vision algorithms in the recent literature designed for other tasks perform on selfie detection as a surrogate task? Can we design and tract the algorithm to be able to scale it for larger and more subtle problems? An attempt has been made in this work to address these questions. In most of the cases, it is observed that humans tend to classify an image as selfie, by noticing the subtle poses of the self-portrait taker. For example, consider the selfie in Fig 2, it can be easily seen that the strong visual cues for inferring a selfie is by making a connection between the shoulder-arm direction (indicated by the red vector) and the head gaze direction (indicated by the blue vector). A short-survey comprising of roughly 50 individuals, from diverse backgrounds, was conducted to verify if other individuals also share a similar opinion. It turned out that most of the people, 36 to be exact, considered hand and head orientation as a factor to classify or group the given images into selfies or non-selfies. Thus, an attempt has been made in this regard and our entire approach consists of three

Word Boundary Estimation for Continuous Speech Using Higher Order Statistical Features

Vijayakrishna Naganoor, Akshay Kumar Jagadish, and Krishnan Chemmangat*
*Department of Electical and Electronics Engineering,
National Institute of Technology Karnataka, India.

***Abstract*—Detection of the start and the end time of words in a continuous speech plays a crucial role in enhancing the accuracy of Automatic Speech Recognition (ASR). Hence, addressing the problem of efficiently demarcating word boundaries is of prime importance. In this paper, we introduce two new acoustic features based on higher order statistics called Density of Voicing (DoV) and Variability of Voicing (VoV) obtained from the bispectral distribution, which when used along with the popular prosodic cues helps in drastically reducing the recognition error rate involved. An ensemble of three different models has been designed to minimize the false alarms, during word boundary detection, by maximizing the uncorrelatedness in prediction from each model. Finally, the majority-voting rule was used to decide if the given frame encompasses a word boundary. The contribution of the work lies in: (i) Converting word boundary detection into a supervised learning problem (ii) Introduction of two new higher order statistical features (iii) Using ensemble methods to find the best model for prediction. Experimental results for NTIMIT Database shows the efficacy of the proposed method and its robustness to noise added during telephonic transmission.**

I. Introduction

Determination of the word boundaries for continuous speech is a challenging task and finds immense significance in the field of Automatic Speech Recognition (ASR) as it helps to reduce the ASR problem into a more simpler single word transcription problem. Word boundary detection can be utilised to extract the Out of Vocabulary (OoV) words such as proper nouns [1] as well as for the rich transcription of speech.

Prosodic features have proven to be superior to word-level information as the speakers use prosody to impose structure on both spontaneous and read speech. Extensive work has been done for German and Indian Language using the prosodic cues [2]. Studies were conducted for observing the behaviour of the pitch pattern across the speech utterances. It was observed that the pitch frequency F_0 fell gradually from the beginning to the end of the utterance. The fact that F_0 rose from the first syllable to the last syllable in a word and fell to the first syllable in the next word was utilized. Log of the average energy [3] is another feature which has been used along with short-term energy [4] to localize the word boundary.

However, these popular acoustic cues often fail to give clues about the word boundaries, particularly when the beginning of a word gets co-articulated with the end of the previous word. The problem becomes further challenging when noise is introduced in the audio files. Basic features become less reliable in the presence of different kinds of sound artifacts and noise, especially when it is non-stationary. In this paper, we propose the usage of the rudimentary acoustic features and higher-order statistical (HOS) features like kurtosis, skewness combined with two new simple yet powerful features, derived from HOS, to improve the robustness of the system. Majority-voting method was used to decide the outcome of each frame from the ensemble of three models namely, Support Vector Machines (SVM), Artificial Neural Network (ANN) and Random Forest Classifier (RFC).

The paper has been organized in the following manner, Section II and III, briefly discusses the features that were explored in the past ([1], [5], [6]) and the additional features that has been proposed in this work. In Section IV, experimental setup is explained with the information about the corpus, classifier setup, evaluation and the implementation of the algorithm. The results are presented in section V. Section VI concludes the work with brief description on future possibilities.

II. Rudimentary Acoustic Features

The following are the basic acoustic cues extracted from the prosodic information and used in word boundary detection:

A. Short-time pitch frequency [4]

Pitch can be defined only for the voiced portion of the speech. It takes on very low values, close to zero, at segments or frames corresponding to the unvoiced region or to those that contain only noise. In general, the frame with word boundary is surrounded by the unvoiced frames and hence, it is in the region where pitch defined is zero.

B. Zero Line Crossing [4]

Zero Line Crossing gives the number of times the the signal crosses the zero mark within the particular frame. This quantity is comparatively lower for frames in and around the voiced portion of speech and has higher values for the segments which correspond to the word to word transition.

C. Log Energy

It has been observed that majority of the frames belonging to the transition between words have lower energy. This could be used as a cue to decide whether the frame is within the a word or closer to a word transition. It is derived from the root-mean squared energy of each frame.

Variable Density Sampling in Compressed Sensing MRI: A Vehicle Routing Approach

Pramit Saha, Satrajit Chakrabarty, Soumya Goswami, Akshay Kumar J, Kasi Rajgopal

***Abstract*—k-space variable density sampling is an effective way of reducing acquisition time in Magnetic Resonance Imaging (MRI). This paper utilizes the notion of compressed sensing by encompassing the variable density sampling problem by a Genetic Algorithm (GA) based Vehicle Routing Problem solver, employed to design an optimal design subspace for perfect reconstruction of sparse signals. Example waveforms are presented for the proposed stochastic drawings, for a variety of design objectives and parameter sets. The methodology has been underpinned by extensive evaluation by various metrics and the results have been compared with several state-of-the-art sampling trajectories.**

***Index Terms*—Compressed sensing, Vehiculear Routing Problem, Variable Density Sampling, *k*-space, Hermitian Symmetry, Time Optimal Smoothing**

I. Introduction

Vehicular Routing Problem, formulated mathematically by Dantzig and Ramser [1] is essentially a NP-hard combinatorial optimization problem. Some superior meta-heuristic algorithms have recently been developed, and Genetic Algorithms (GA) have been shown to be capable of solving VRPs [2].

Compressed Sensing [3] in MRI is a theoretical framework which demonstrates the design of a measurement ensemble for accurate reconstruction of incoherent sparse signals, from a limited number of linear sub-Nyquist projections through non-linear and iterative algorithms. CS justifies the possibility of recovering an image from a significantly undersampled Fourier measurement domain of the MR image to overcome the coherence barrier [4] which has culminated in the popularity of k-space undersampling using VDS. As most of the energy of an image is concentrated around the k-space center in the low-frequency region [3], a VDS may adequately sample the central k-space region to reduce low-frequency aliasing artifacts and undersample the outer k-space region to reduce scan time and to increase resolution.

In a recent work[5], Chauffert develops an optimal sampling distribution based on independent drawings, in *k*-space with improved image sparsity and sample rejection for any re-visited position, that guarantees better reconstruction. The authors report that the reconstruction quality with heuristic sampling strategies is much better than the theoretical ones owing to the structured image sparsity. The key features of VDS viz. the limit of empirical measures and their mixing properties are elaborated in [6], where the authors have strongly justified the cause of deterministic sampling around the *k*-space centre. In [5], a local random walk sampling based on Markov chains has been presented, enforcing the samples to fit continuous trajectories. But the continuity is strongly dependent on the parameter α in a way that an increase in problem dimension or decrease in continuity offers considerable hardware limitations in MRI, where as an increase in continuity may deteriorate the reconstruction quality. Another continuous trajectory based on global travelling salesman problem (TSP) solver has been employed in [7], that emulates a VDS strategy, but the problem of gradient constraints on MRI scanner still persists. This very problem has been solved in [8], by projecting it onto the set of hardware constraints using convex optimisation, that yields faster trajectories and gives better reconstruction quality at a given scan time.

A closer look reveals certain drawbacks in the TSP strategy. As a single salesman is engaged to traverse all the cities, the movement becomes very chaotic and insignificant without preserving any directional characterisation. Furthermore, the route trajectory imposed on randomly distributed cities leaves the k-space centre rarely sampled, deteriorating reconstruction image quality. Hence, there is an inherent urge of extending the TSP solver to an ameliorated approach employing an optimal number of salesmen that may solve these issues. Moreover, incorporating multiple salesman in the form of vehicle routing problem executes the traversing operations parallelly all over the k space, enunciating a faster convergence. Since parallel processing or routing is now possible, there is a drastic reduction in acquisition time that serves as a great advantage. As the GA based heuristic solution of the path itself describes a piece-wise linear continuous trajectory, it has been utilized here for semi-perfect sparse MRI reconstruction by reaching a near-optimal solution in minimum number of iterations. An overview of the proposed methodology has been demonstrated with the help of a block diagram in Figure 2.

II. Vehicular Routing Problem

A. Theory

The Vehicle Routing Problem (VRP) can be described as the problem of designing optimal delivery or collection routes from one or several depots to a number of geographically dispersed destinations (cities or customers) with specific

P. Saha is with the Department of Electrical Engineering, Jadavpur University, Kolkata, India, e-mail: pramitsaha.juee@gmail.com

S. Chakrabarty is with the Department of Electrical Engineering, Jadavpur University, Kolkata, India, e-mail: satrajit2012@gmail.com

S. Goswami is with the Department of Electrical Engineering, Jadavpur University, Kolkata, India, e-mail: soumyagoswami.juee@gmail.com

A. Kumar J is with the Department of Electrical and Electronics Engineering, National Institute of Technology Karnataka, Surathkal, India, e-mail: akshaykumar3495@gmail.com

K. Rajgopal is with the Department of Electrical Engineering, Indian Institute of Science, Bangalore, India, e-mail: kasi@ee.iisc.ernet.in