Physical Sciences Research International
Vol. 1(2), pp. 32-35, May 2013
Short Communication

Distance measure with regular look-up for sub-sequence differences decreased method

Mahdi Jalali

Department of Electrical Engineering, Naghadeh Branch, Islamic Azad University, Naghadeh, Iran.

Accepted 26 April, 2013

ABSTRACT

In this paper, we calculated the absolute value of the differences in this sub-sequence value changes. We further calculated the average of these differences. We would expect the average of the differences to be monotonically decreasing as the signal changes during the estimation process. This average of differences decreased monotonically in all iterations but, the last outer loop iteration. Algorithms imply the waveforms of the estimated clean speech from the noisy speech signals with 5 and 0 dB SNR.

Keywords: Table look-up, STFT, fast calculation, norm, subsequence method.

E-mail: jalali.mahdi@gmail.com.

INTRODUCTION

For the case that a special table is created similar to the regular look-up table for Term 3, with the exception that the prior joint probabilities are estimated from only one clique. In this case, the value of was estimated from the values of the regular look-up table for Term 1, the regular look-up table for Term 2 and the special look-up table for Term 3. Similarly, a special table is created for Term 3 for the case in which the prior joint probabilities these four special tables are all 5-dimensional tables of type short. The band number is not an index to these tables. The indices are: one index for the noisy grayscale bin number, and four indices for the binary configuration of the clique of interest, excluding the value. We define a distance measure on the norms of the square roots of the spectrograms, or in other words on the norms of STFT images, rather than on the spectrograms directly (Demange et al., 2006). In Roman et al. (2003); Agnew et al. (2000); Carvalho et al. (2003) and Jalali et al. (2012) an iterative algorithm for searching for such a signal is described. That algorithm assumes that the modified STFT image of the whole speech signal is available at the time

METHODOLOGY

In each Metropolis iteration, the value must be calculated in order to decide whether a certain pixel value should be changed. This is the reason that running such an iterative algorithm is so time consuming, the calculation is expensive and must be preformed frequently. For that reason, we attempted to pre-calculate and store in advance look-up tables that facilitate the calculation as much as possible. Also recalled is the product of the band-dependent posterior probabilities over the pixels. That means that the band-dependent posterior probabilities are equal for all pixels, except for pixels (which we define to be the pixel above image whose clean version we are trying to estimate. The value can be re-written as the product of the following three ratios of posterior probabilities. We refer to these three ratios, without the power, as Term 1, Term 2 and Term 3, respectively. Notice that the three terms in the calculation are similar but not equivalent. Term 1 describes the change in the posterior probability of pixel when the value is changed. This is clearly different from Term 2, which describes the change in the posterior probability of pixel when its value is changed. Term 3 is similar to Term 1 but describes the change of the posterior probability of pixel when the color of ˆhis changed. Similarly, only two terms exist in the calculation and in Term 2 the prior join probabilities of ˆ in image are estimated from only one clique. Also notice that if ˆhis in row, three terms exist in the calculation but in Term 3 the prior joint probabilities images are estimated from only one clique. Similarly, three terms exist in the calculation but in Term 1 the prior joint probabilities of images are estimated from only one clique. To simplify the calculation of the value performed in the Metropolis Algorithm, several look-up tables were created and stored in advance. It is a description of these look-up tables and one more look-up table created for later needs, which is described first. One table is of type float (4 bytes) and contains the estimated band-dependent posterior probability. This

Table 1. A table summarizing the timings of the different steps.

Event	Time (s)	Note
Create training set:		Performed once
Pre-processed signals	4.440	For all 5 data-sets
Create noise signal	0.289	For each SNR level
Create noisy speech signals	10.990	For all 5 data-sets and both SNR levels
Create clean spectrogram	0.101	For each of the 5 data-sets
Create clean binary image	2.051	For all 4 training data-sets
Estimate models:		Performed once
Estimate noise model	0.933	For each SNR level
Estimate binary model and posterior probability	0.165	For each SNR level
Estimate clean speech:		Performed for each testing signal
Create noisy spectrogram	0.101	
Estimate hard and soft segmentations	236.555	
Estimate clean spectrogram	0.004	
Estimate clean speech	0.225	

table is an 8-dimensional table whose indices are: one index for the band number, one index for the noisy grayscale bin number bin, and six indices for the binary neighborhood configuration Three look-up tables were created, each one relating to one of the three terms in the calculation. These three look-up tables, which we refer to as the regular look up tables, are of type short. One of table is created for Term 2 and contains the value which approximates the Term 2. The value is chosen to be a large positive integer for which the value.

RESULTS AND DISCUSSION

The timings reported here were performed on a Intel Centrino Duo, 1.66 GHz each, 2 GB of RAM and running on Windows. The programs were compiled with GNU gcc (version 4.1.2), with the highest level of optimization. We report on the user CPU time as reported by the Windows “time” command. Table 1 displays a summary of the timings described in more detail below. The time to create our training sets from the clean signals, which needs to be only performed once, is as follows: It took 4.440 s to pre-process the signals in all 5 data-sets. It took an additional 2.459 s to add the 500 ms of silence to the beginning and end of each of the pre-processed signals.

It took 0.289 s to create each of the two noise signals, one for each SNR level. It took 10.990 s to create the noisy speech signals by adding the clean speech signals to the noise signal segments for both SNR levels and all 5 data-sets. It took 0.101 s to create the spectrogram (and simultaneously the spectrogram image) of each clean speech signal. The clean binary images were created by thresholding the clean spectrogram images in the four training data-sets, which took 2.051 s. This concludes the steps required to creating a training set. It took 0.933 s to estimate the noisy information probability for each SNR level from the training set. It took 0.165 s to estimate the posterior probability for each SNR level from the training set. It took 0.101 s to create the spectrogram of each noisy speech signal. It took 236.555 s to estimate simultaneously the hard segmentation and soft segmentation of each noisy speech spectrogram. This is clearly the bottleneck of our methodology; it takes 3 orders of magnitude longer than any other step in our methodology. It took 0.004 s to estimate the clean spectrogram using the soft segmentation and the noisy spectrogram. It took 0.225 s to estimate the clean speech signal from the estimated clean spectrogram. Notice that these are timings for processing one whole signal on the order of about 1.5 s. For testing purposes, the estimated clean speech signals were post-processed. It took 0.944 s to post-process all the testing signals estimated by our methodology, an equivalent amount of time to post-process all the testing signals estimated by the gold standard methodology, and an equivalent amount of time to post-process all the noisy testing signals. For display purposes, the spectrogram images of the post-processed estimated clean speech signals by our methodology and the gold standard methodology were created, as well as the spectrogram images of the post-processed noisy testing signals. It took 0.101 s to create the spectrogram and spectrogram image of each signal.

Conclusion

We then created the spectrogram of the estimate speech. The original clean spectrogram and the spectrogram of the estimated speech were indeed very close. In fact, the spectrogram images were identical grayscale bin number, and six indices for the binary neighborhood configuration that excludes the value is the other two regular look-up tables were pre-calculated in the same way for Terms. For Term contain the values look-up table for Term 2 and regular look-up table for Term 1.

REFERENCES

Agnew J, Thornton JM, 2000. Just noticeable and objectionable group delays in digital hearing aids. J Am Acad Audiol, 11:330–336.

Carvalho BM, Herman GT, Matej S, Salzberg C, Vardi E, 2003. Binary to mography for triplane cardiography. In: Information Processing in Medical Imaging, Kuba A, Samal M and Todd-Pokropek A (Eds.), Springer-Verlag, pp: 29–41.

Demange S, Cerisara C, Haton JP, 2006. Mask estimation from missing data Recognition using back ground noise sniffing. Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing, May.

Jalali M, Sedghi T, 2011. Semi Supervised Feature Extraction for Filling Semantic Gap in Image Retrieval. In Proc. of IEEE, Machine Vision and Image Processing Symposium, Iran.

Roman N, Wang D, Brown GJ, 2003. Speech segregate on based on sound localization. J Acoustical Soc Am, 114: 2236–2252.