ROBUST AUTOMATIC PHONEME RECOGNITION FEATURES USING COMPLEX WAVELET PACKET TRANSFORM COEFFICIENTS

Tjong Wan Sen [#1]

[#] *Departemen Sistem Komputer*

Institut Teknologi Harapan Bangsa

Jl. Dipati Ukur 80-84, Bandung

[1]wansen@ithb.ac.id

***Abstrak*— Untuk meningkatkan kinerja sistem pengenalan fonem otomatis pada saat dioperasikan pada lingkungan berderau, kami mengembangkan teknik baru yang dapat melakukan estimasi terhadap suatu fitur fonem bersih dari bentuk berderaunya. Fitur-fitur kokoh tersebut diperoleh dari koefisien transformasi paket wavelet kompleks (Complex Wavelet Packet Transform/CWPT). Karena koefisien CWPT merepresentasikan semua pita frekuensi yang berbeda dari suatu sinyal masukan, mendekomposisi sinyal masukan tersebut ke dalam pohon CWPT yang lengkap akan mencakup semua frekuensi yang terlibat dalam proses pengenalan. Setiap komponen frekuensi dalam sinyal masukan akan ditempatkan pada tepat satu pita frekuensi yang spesifik. Untuk suatu campuran sinyal domain waktu dengan frekuensi yang berbeda-beda, misalnya sinyal fonem dengan derau, semua koefisien fonem dalam pita frekuensi yang sama, yaitu semua koefisien yang melewati jalur filter bank wavelet yang sama, akan berubah sesuai dengan magnituda komponen frekuensi derau. Oleh karena itu, jika ada sebuah pita frekuensi yang tidak mengandung derau sama sekali, seluruh koefisien fonem pada pita frekuensi tersebut tidak akan mengalami perubahan. Informasi dari semua koefisien yang dikandung oleh pita frekuensi tersebut kemudian dapat dimanfaatkan untuk melakukan estimasi terhadap kemungkinan fonem bersihnya. Karena jumlah fonem dalam suatu bahasa adalah terbatas dan relatif kecil dan sudah diketahui dengan baik sebelumnya, teknik yang dikembangkan ini fisibel secara komputasi. Hasil-hasil simulasi menunjukkan bahwa teknik baru yang dikembangkan ini merupakan pengekstrak fitur yang efisien dan tidak hanya dapat meningkatkan kekokohan sistem pengenal fonem otomatis jika dioperasikan pada berbagai macam lingkungan yang berderau tetapi juga tetap memelihara kinerja baiknya pada lingkungan yang bersih.**

***Keywords*— pengenalan fonem otomatis, kokoh, derau, transformasi paket wavelet kompleks.**

***Abstract*— To improve the performance of Automatic Phoneme Recognition in noisy environment, we developed a new technique that could estimate clean phoneme feature from its noisy one. These robust features are obtained from Complex Wavelet Packet Transform (CWPT) coefficients. Since the CWPT coefficients represent all different frequency bands of the input signal, decomposing the input signal into complete CWPT tree would covered all frequencies that involved in recognition process. Each frequency would be placed into exactly one of its frequency bands. For time overlapping signals with different frequency contents, e. g. phoneme with noises, all coefficients belongs to the same frequency band, which is coming through the same wavelet filter banks path, would be changed according to noise frequencies magnitude. Thus, if there is one frequency band which contain no noises at all, all coefficients belongs to that frequency band would not change. Information from all coefficients belongs to that frequency band could be used then to estimate the clean phonemes. Since the numbers of phonemes are limited and already well known, this technique is computationally feasible. Simulation results showed that this new technique is an efficient features extractor that improves the robustness of the systems in various adverse noisy conditions but still reserve the good performance in clean environments.**

***Keywords*— phoneme recognition, robust feature, complex wavelet packet, coefficients, noise.**

I. INTRODUCTION

This document is a template. An electronic copy can be downloaded from the conference website. For questions on paper guidelines, please contact the conference publications committee as indicated on the conference website. Information about final paper submission is available from the conference website.

II. PAGE LAYOUT

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

A. Page Layout

Your paper must use a page size corresponding to A4 which is 210mm (8.27") wide and 297mm (11.69") long. The margins must be set as follows:

- Top = 19mm (0.75")

- Bottom = 43mm (1.69")
- Left = Right = 14.32mm (0.56")

Your paper must be in two column format with a space of 4.22mm (0.17") between columns.

III. Page Style

All paragraphs must be indented. All paragraphs must be justified, i.e. both left-justified and right-justified.

A. *Text Font of Entire Document*

The entire document should be in Times New Roman or Times font. Type 3 fonts must not be used. Other font types may be used if needed for special purposes.

Recommended font sizes are shown in Table 1.

B. *Title and Author Details*

Title must be in 24 pt Regular font. Author name must be in 11 pt Regular font. Author affiliation must be in 10 pt Italic. Email address must be in 9 pt Courier Regular font.

TABLE I
FONT SIZES FOR PAPERS

Font Size	Appearance (in Time New Roman or Times)		
	Regular	**Bold**	**Italic**
8	table caption (in Small Caps), figure caption, reference item		reference item (partial)
9	author email address (in Courier), cell in a table	abstract body	abstract heading (also in Bold)
10	level-1 heading (in Small Caps), paragraph		level-2 heading, level-3 heading, author affiliation
11	author name		
24	title		

All title and author details must be in single-column format and must be centered.

Every word in a title must be capitalized except for short minor words such as “a”, “an”, “and”, “as”, “at”, “by”, “for”, “from”, “if”, “in”, “into”, “on”, “or”, “of”, “the”, “to”, “with”.

Author details must not show any professional title (e.g. Managing Director), any academic title (e.g. Dr.) or any membership of any professional organization (e.g. Senior Member IEEE).

To avoid confusion, the family name must be written as the last part of each author name (e.g. John A.K. Smith).

Each affiliation must include, at the very least, the name of the company and the name of the country where the author is based (e.g. Causal Productions Pty Ltd, Australia).

Email address is compulsory for the corresponding author.

C. *Section Headings*

No more than 3 levels of headings should be used. All headings must be in 10pt font. Every word in a heading must be capitalized except for short minor words as listed in Section III-B.

1) Level-1 Heading: A level-1 heading must be in Small Caps, centered and numbered using uppercase Roman numerals. For example, see heading “III. Page Style” of this document. The two level-1 headings which must not be numbered are “Acknowledgment” and “References”.

2) Level-2 Heading: A level-2 heading must be in Italic, left-justified and numbered using an uppercase alphabetic letter followed by a period. For example, see heading “C. Section Headings” above.

3) *Level-3 Heading:* A level-3 heading must be indented, in Italic and numbered with an Arabic numeral followed by a right parenthesis. The level-3 heading must end with a colon. The body of the level-3 section immediately follows the level-3 heading in the same paragraph. For example, this paragraph begins with a level-3 heading.

D. *Figures and Tables*

Figures and tables must be centered in the column. Large figures and tables may span across both columns. Any table or figure that takes up more than 1 column width must be positioned either at the top or at the bottom of the page.

Graphics may be full color. All colors will be retained on the CDROM. Graphics must not use stipple fill patterns because they may not be reproduced properly. Please use only *SOLID FILL* colors which contrast well both on screen and on a black-and-white hardcopy, as shown in Fig. 1.

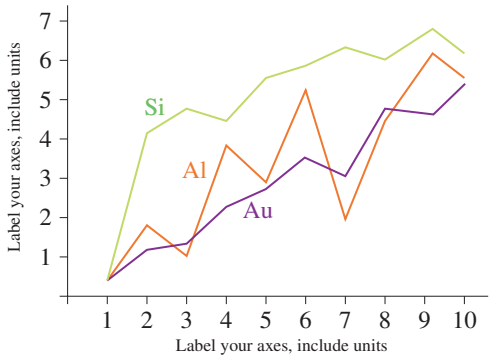

Fig. 1 A sample line graph using colors which contrast well both on screen and on a black-and-white hardcopy

Fig. 2 shows an example of a low-resolution image which would not be acceptable, whereas Fig. 3 shows an example of an image with adequate resolution. Check that the resolution is adequate to reveal the important detail in the figure.

Please check all figures in your paper both on screen and on a black-and-white hardcopy. When you check your paper on a black-and-white hardcopy, please ensure that:

- the colors used in each figure contrast well,
- the image used in each figure is clear,
- all text labels in each figure are legible.

E. *Figure Captions*

Figures must be numbered using Arabic numerals. Figure captions must be in 8 pt Regular font. Captions of a single

line (e.g. Fig. 2) must be centered whereas multi-line captions must be justified (e.g. Fig. 1). Captions with figure numbers must be placed after their associated figures, as shown in Fig. 1.

Fig. 2 Example of an unacceptable low-resolution image

Fig. 3 Example of an image with acceptable resolution

F. Table Captions

Tables must be numbered using uppercase Roman numerals. Table captions must be centred and in 8 pt Regular font with Small Caps. Every word in a table caption must be capitalized except for short minor words as listed in Section III-B. Captions with table numbers must be placed before their associated tables, as shown in Table 1.

G. Page Numbers, Headers and Footers

Page numbers, headers and footers must not be used.

H. Links and Bookmarks

All hypertext links and section bookmarks will be removed from papers during the processing of papers for publication. If you need to refer to an Internet email address or URL in your paper, you must type out the address or URL fully in Regular font.

I. References

The heading of the References section must not be numbered. All reference items must be in 8 pt font. Please use Regular and Italic styles to distinguish different fields as shown in the References section. Number the reference items consecutively in square brackets (e.g. [1]).

When referring to a reference item, please simply use the reference number, as in [2]. Do not use "Ref. [3]" or "Reference [3]" except at the beginning of a sentence, e.g. "Reference [3] shows …". Multiple references are each numbered with separate brackets (e.g. [2], [3], [4]–[6]).

Examples of reference items of different categories shown in the References section include:

- example of a book in [1]
- example of a book in a series in [2]
- example of a journal article in [3]
- example of a conference paper in [4]
- example of a patent in [5]
- example of a website in [6]
- example of a web page in [7]
- example of a databook as a manual in [8]
- example of a datasheet in [9]
- example of a master's thesis in [10]
- example of a technical report in [11]
- example of a standard in [12]

IV. Conclusions

The version of this template is V2. Most of the formatting instructions in this document have been compiled by Causal Productions from the IEEE LaTeX style files. Causal Productions offers both A4 templates and US Letter templates for LaTeX and Microsoft Word. The LaTeX templates depend on the official IEEEtran.cls and IEEEtran.bst files, whereas the Microsoft Word templates are self-contained. Causal Productions has used its best efforts to ensure that the templates have the same appearance.

Acknowledgment

The heading of the Acknowledgment section and the References section must not be numbered.

Causal Productions wishes to acknowledge Michael Shell and other contributors for developing and maintaining the IEEE LaTeX style files which have been used in the preparation of this template. To see the list of contributors, please refer to the top of file IEEETran.cls in the IEEE LaTeX distribution.

References

[1] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[3] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.

[4] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, "High resolution fiber distributed measurements with coherent OFDR," in *Proc. ECOC'00*, 2000, paper 11.3.4, p. 109.

[5] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.

[6] (2002) The IEEE website. [Online]. Available: http://www.ieee.org/

[7] M. Shell. (2002) IEEEtran homepage on CTAN. [Online]. Available: http://www.ctan.org/tex-archive/macros/latex/contrib/supported/IEEEtran/

[8] *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[9] "PDCA12-70 data sheet," Opto Speed SA, Mezzovico, Switzerland.

[10] A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.

[11] J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of TCP Reno congestion avoidance and control," Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.

[12] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.