

## 本期要目

壹、ROCLING 2010 CALL FOR PAPERS	第二頁
貳、CFP- IEEE Transactions on Affective Computing	第三頁
參、專文-現階段語音文件摘要研究之簡介(林士翔、陳柏琳)	第四~十六頁

### 獎助學生出席國際會議

#### 獎助會議：

- |              |           |
|--------------|-----------|
| 1. COLING    | 2. ACL    |
| 3. ACM SIGIR | 4. ICASSP |

#### 獎助說明：

- 申請人須同時具備下列資格：
  - 被接受論文之第一作者(指導教授不計)。
  - 本會會員。
  - 投稿時為國內在學學生。
- 獎助金額：由審查委員會依地區別及論文等級審定獎助金額，每名獎助金額上限為美金1,000元。
- 獎助名額：每個會議獎助一~二名。

#### 申請辦法：

- 申請期限：論文被接受發佈日起兩週內提出。
- 申請手續：申請人需將論文接受函、審查意見、學生證、論文全文及申請書等相關資料郵寄至本會秘書處。(申請書請至<http://www.aclclp.org.tw/doc/fundreg.htm> 下載)

#### 受獎助人義務：

- 出席會議發表論文。
- 論文全文必須以書面同意投稿至本會期刊。
- 代學會攜去宣傳品及帶回相關資料。

### IJCLCLP - Call for Papers

#### Special Issue on Web 2.0 and Social Computing

With the popularity of Web 2.0 in recent years, there are an increasing number of applications targeting at not only content but people and their social connections. People interact with each others throughout social networking media such as blogs, Wikis, Flickr, Twitter, and Del.icio.us. Consequently, technologies developed for natural language processing have been widely applied on analyzing and extracting social media. This special issue brings together the computational linguistics with Web 2.0 and social computing. Through soliciting the state-of-the-art research methods and results on the relevant topics, this special issue aims at discovering synergies among the computational linguistics, Web 2.0 and social computing in order to identify new ways of understanding this emerging field.

#### Areas of interest include, but are not limited to:

- Natural Language Applications on Web 2.0 and Social Networks
- Social Network Analysis
- Search and Mining on Web 2.0 and Social Media

#### Schedule

Submission deadline:	July 1, 2010
Notification of acceptance:	September 1, 2010
Final manuscript due:	November 1, 2010
Tentative publication date:	December 15, 2010

For more details, please visit our website:  
<http://www.aclclp.org.tw/journal/cfp.php>

# Call for Papers

## Conference on Computational Linguistics and Speech Processing

### 第二十二屆自然語言與語音處理研討會, ROCLING XXII

National Chi Nan University, Nantou, Taiwan

1- 2 September 2010

The 22<sup>nd</sup> ROCLING Conference will be held at National Chi Nan University, Nantou, on September 1-2, 2010. Sponsored by Association for Computational Linguistics and Chinese Language Processing (ACLCLP), ROCLING is the most historied and major conference in the broad field of computational linguistics, speech processing, and related areas in Taiwan.

ROCLING XXII will be co-hosted by the Department of Computer Science and Information Engineering and the Department of Electrical Engineering, National Chi Nan University. The two-day conference will feature invited talks, paper, and poster sessions.

ROCLING XXII invites submissions of original and unpublished research papers on all areas of computational linguistics, natural language processing, and speech processing, including, but not limited to, the following topic areas.

#### Topics of Interest:

- cognitive/psychological linguistics
- discourse/dialogue modeling
- information extraction/text mining
- information retrieval
- language understanding/generation
- lexicon/morphology
- machine translation/multilingual processing
- named entity recognition
- NLP applications/tools/resources
- phonetics/phonology
- question answering
- semantic web
- semantics/pragmatics
- speech analysis/synthesis
- speech recognition/understanding
- spoken dialog systems
- spoken language processing
- syntax/parsing
- text/speech summarization
- web knowledge discovery
- word segmentation/POS tagging
- others

#### Important Dates:

Preliminary paper submission deadline:	July 2, 2010
Notification of acceptance:	August 6, 2010
Camera-ready due:	August 15, 2010
Conference date:	September 1-2, 2010



The new *IEEE Transactions on Affective Computing* seeks original manuscripts for publication. This new online only journal will publish cross disciplinary and international archival results of research on the design of systems that can recognize, interpret, and simulate human emotions and related affective phenomena. The journal will publish original research on the principles and theories explaining why and how affective factors condition interaction between humans and technology, on how affective sensing and simulation techniques can inform our understanding of human affective processes, and on the design, implementation and evaluation of systems that carefully consider affect among the factors that influence their usability. Surveys of existing work will be considered for publication when they propose a new viewpoint on the history and the perspective on this domain. The journal covers but is not limited to the following topics:

#### Sensing & Analysis

- Algorithms and features for the recognition of affective state from face and body gestures
- Analysis of text and spoken language for emotion recognition
- Analysis of prosody and voice quality of affective speech
- Recognition of auditory and visual affect bursts
- Recognition of affective state from central (e.g. fMRI, EEG) and peripheral (e.g. GSR) physiological measures
- Methods for multi-modal recognition of affective state
- Recognition of group emotion
- Methods of data collection with respect to psychological issues as mood induction and elicitation or technical methodology as motion capturing
- Tools and methods of annotation for provision of emotional corpora

#### (Cyber)Psychology & Behavior

- Clarification of concepts related to 'affective computing' (e.g., emotion, mood, personality, attitude) in ways that facilitate their use in computing.
- Computational models of human emotion processes (e.g., decision-making models that account for the influence of emotion; predictive models of user emotional state)
- Studies on cross-cultural, group and cross-language differences in emotional expression
- Contributions to standards and markup language for affective computing

#### Behavior Generation & User Interaction

- Computational models of visual, acoustic and textual emotional expression for synthetic and robotic agents

- Models of verbal and nonverbal expression of various forms of affect that facilitate machine implementation
- Methods to adapt interaction with technology to the affective state of users
- Computational methods for influencing the emotional state of people
- New methods for defining and evaluating the usability of affective systems and the role of affect in usability
- Methods of emotional profiling and adaptation in mid- to long-term interaction
- Application of affective computing including education, health care, entertainment, customer service, design, vehicle operation, social agents/robotics, affective ambient intelligence, customer experience measurement, multimedia retrieval, surveillance systems, biometrics, music retrieval and generation

Editor in Chief: Jonathan Gratch, University of Southern California. To submit a paper, go to <https://mc.manuscriptcentral.com/taffc-cs>.



Published in cooperation with: IEEE Robotics and Automation Society, IEEE Signal Processing Society, IEEE Society on Social Implications of Technology, IEEE Consumer Electronics Society, IEEE Circuits and Systems Society

# 現階段語音文件摘要研究之簡介

林士翔、陳柏琳

國立台灣師範大學資訊工程系

{shlin, berlin}@csie.ntnu.edu.tw

## 一、前言

在現今資訊科技與網際網路快速蓬勃發展的時代，大量的文字或多媒體影音資訊被快速地傳遞與分享於全球各地，資訊超載(Information Overload)問題也因此隨之產生。如何能讓人們快速地、有效率地瀏覽與日俱增的文字資訊或多媒體影音資訊，已是一個刻不容緩的研究課題。而在眾多的研究方法中，文件摘要(Document Summarization)被視為是一項不可或缺的關鍵技術[1-4]。

文件摘要旨在於擷取單一文件(Single-Document)或多文件(Multi-Document)中內含的重要語意與主題資訊，並以適當的方式呈現，讓使用者可以快速地瀏覽與理解文件中的重要資訊。因此，使用者不需花費大量的時間與力氣逐一地審視或過濾文件內容，便可以快速獲得其所需的資訊。另一方面，文件摘要技術對於如何能自動地、有效率地處理具時序性的多媒體影音內容，諸如電視及廣播新聞、語音郵件、會議及演講錄音等，更是顯得非常重要[3]。其原因在於對於動輒數分鐘到數小時的多媒體影音，使用者通常無法如瀏覽文字文件般的任意(或是隨機)跳躍地瀏覽文件中的任何一個部份，而必須耐心地從頭到尾閱讀或聆聽整份多媒體影音內容，才能理解其中所描述的語意與主題。雖然對於含有語音訊號的多媒體影音，我們可透過自動語音辨識(Automatic Speech Recognition, ASR)的技術將其轉換成易於瀏覽的文字內容，再經由文字文件摘要(Text Document Summarization)技術的處理，達到摘要多媒體影音或者其他語音文件(Spoken Documents)的目的。但就現階段語音辨識技術的發展而言，語音文件經語音辨識自動產生的文字轉寫結果，仍存有一定程度的辨識錯誤，並且缺乏像章節與標點符號等結構資訊；再加上語音文件本身所包含的許多口語助詞、遲疑、重覆等內容，使得語音文件摘要(Spoken Document Summarization)技術的發展面臨更多的挑戰。所以，近年來語音文件摘要技術的發展在語音及自然語言處理領域已是一個愈來愈受重視的研究課題[2-3]。

由於目前仍有許多語音文件摘要研究所使用的摘要技術是源自於傳統文字文件摘要研究的成果，因此本文將先簡略地回顧過去在文字文件摘要上一些重要的相關研究與成果；然後，闡述當前語音文件摘要研究的發展現況，並介紹數個文件摘要效能的評估方式；最後是結論與對於語音文件摘要研究的未來展望。

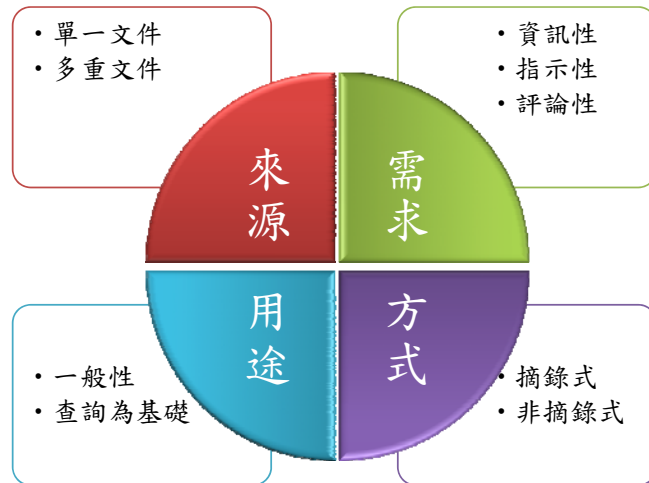


圖 1. 自動摘要技術分類圖

## 二、文字文件摘要技術

一般而言，文件摘要技術可從許多不同構面進行探討，包括來源、需求、用途及方式，分類圖如圖 1 所示，下列將簡述各個不同構面的差異[1]：

1. 來源：根據文件來源，可以分為單一文件摘要與多文件摘要；而在多文件摘要中，尤其需額外考慮文件間彼此的資訊重複(Redundancy)性[5]與文件所描述事件的發生先後順序(Causality)[6]。
2. 需求：依據使用者需求不同，摘要內容可以是具有資訊性(Informative)，用來表達文件本身所含的重要語意及主題資訊；或者具指示性(Indicative)，可提供如文件分類等文件處理之依據；亦或具評論性(Critical)，分別將文件正面及反面觀點(Positive and Negative Sentiments)[7]呈現出來。
3. 方式：可概分為二大類，摘錄式(Extractive)摘要與非摘錄式(Abstractive)摘要(或重寫摘要)。前者主要是依特定摘要比例，從原文件中選出重要的文句、段落或章節，經適當的修飾與串接而組成摘要；而後者則是根據文件內容自動地重新產生一段簡短扼要的摘要來代表文件的語意及主題資訊，其中所使用之詞彙或慣用語不一定是全然地來自於原始文件，是一種較為貼近人們日常撰寫摘要的方式。
4. 用途：依摘要用途可分為一般性(Generic)摘要與以查詢為基礎(Query-based)的摘要，前者呈現文件全面性的主題，摘要內容以涵蓋整篇文件所有重要主題為主；後者則根據使用者或特定的查詢來呈現文件中與查詢相關的摘要資訊。

然而因為非摘錄式摘要需要較多複雜的自然語言處理(Natural Language Processing, NLP)技術，如資訊擷取(Information Extraction)及對話理解(Discourse Understanding)及自然語言生成(Natural Language Generation)等技術[8-9]，因此近年來研究主要還是以摘錄式摘要為主。自動摘要技術發展概況如圖 2 所示，以下將就摘錄式摘要的研究與進展作簡短回顧。

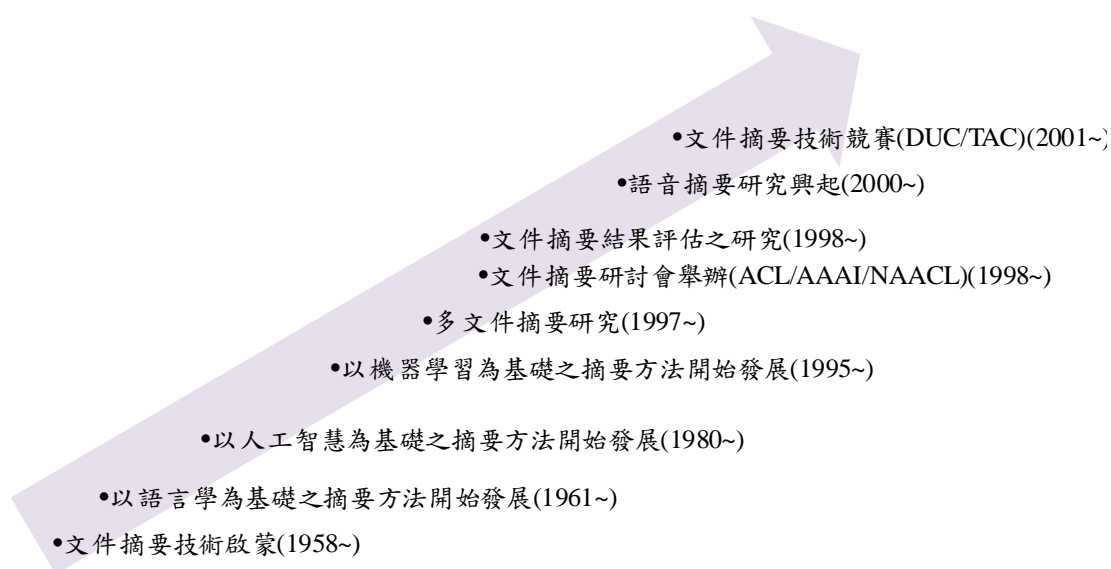


圖 2. 自動摘要技術發展概況

傳統文字文件摘要技術發展最早可追溯至 1950 年代末期，IBM 學者 Luhn 提出以使用詞頻(Frequency)來評量詞的重要性與計算文件中每個語句的顯著性(Significance Factor)[10]，再將語句依其顯著分數進行排序(由高至低)，最後根據特定摘要比例進行摘錄式摘要的產生。此研究不僅自此開啓了自動文件摘要研究的大門，同時亦奠定了許多日後研究可以依循的準則，例如每個詞彙可以進行詞幹分析(Stemming)將其還原成詞根(Root Form)、移除停用詞(Stop Word)的影響，及計算實詞(Content Word)的重要性等。隨後，學者 Baxendale 在審視 200 篇科技文章後，發現有 85%的重要語句會出現在文章中的第一段、以及 7%的重要語句會出現在最後一段[11]。因此，提出語句的在文章中的位置(Position)資訊是在進行文件摘要的重要語句選取時的一項關鍵線索。常常被拿來定義使用。再者，學者 Edmundson[12]除了使用詞頻與位置資訊外，並提出可以額外地使用線索慣用語(Cue Phrases)與詞彙出現在文件不同部份(Skeleton)(如標題、前言、結論等位置)的頻率特徵，再結合統計式方法，對上述這些種特徵作線性組合以計算語句重要性。經過上述三個先探研究(Pilot Studies)後，文件摘要便逐漸成為自然語言處理的一項重要的研究課題。以下，本文將過去研究所陸續發展出的文件摘要方法概略地區分為二大類 [1]：

#### 1. 以語段為基礎(Discourse-based Approach)之摘要方法

此類方法通常是利用自然語言分析(Natural Language Analysis)技術對文章結構進行剖析，接著再根據文法結構(Grammar Structure)與語言機制(Linguistic Devices)，如

首語重複(Anaphora)、省略(Ellipsis)、結合(Conjunction)，或同義詞(Synonymy)、上義詞(Hypernym)等詞彙關係(Lexical Relation)決定不同語段的凝聚(Cohesion)關係，以進行自動摘要。此類方法較著名的相關研究包括使用詞彙鏈(Lexical Chain) [13]、宏觀語段結構(Discourse Macro Structure)[14]、修辭結構(Rhetorical Structure)[15-17]等。

## 2. 以機器學習為基礎(Machine-learning-based Approach)之摘要方法

此類的研究可追溯至於 1990 年代，因機器學習(Machine Learning)技術逐漸被介紹至自然語言處理而興起。一般來說可分為兩個面向：非監督式(Unsupervised)摘要模型與監督式(Supervised)摘要模型的使用。非監督式模型通常產生單一種摘要特徵供語句排序使用，例如語句與文章相關性[18]、語句所形成的語言模型生成文件之機率等[19]、語句間關係性[20-21]、語句中詞彙在機率式潛藏主題空間(Probabilistic Latent Topic Space)的重要性[22]、或語句與文件在機率式潛藏主題空間的模型距離[23-25]等，這些作法基本上在建立摘要模型時並不需參照或使用任何人工事先標記過的語料庫。另一方面，監督式模型則通常需要有人工事先標記過的語料庫供模型訓練使用，亦即會使用到一組訓練文件集以及其中每篇訓練文件的每一語句是否屬於摘要語句(Summary Sentence)的資訊來做為模型訓練之基礎。通常，監督式模型可同時結合多種摘要特徵來表示每一語句(其中的每一種特徵可以是上述以語段為基礎摘要方法或是非監督式摘要模型針對此語句所輸出的分數或機率值)，這些特徵合起來所形成的特徵向量可來用來做為監督式摘要模型判斷此語句是否屬於摘要語句的依據[26]。此類方法較著名的相關研究包括簡單貝氏分類器(Naïve-Bayes Classifier)[27]、高斯混合模型(Gaussian Mixture Model, GMM)[28]、隱藏式馬可夫模型(Hidden Markov Model, HMM)[29]、支援向量機(Support Vector Machine, SVM)[30]、條件隨機場域(Conditional Random Fields, CRF)[31-32]等。

## 三、語音文件摘要技術

目前多媒體影音已經佔據了網際網路一半以上的流通量，雖然透過視覺線索可以對多媒體影音內涵進行瀏覽與搜尋，但其所含的語音文件卻能提供更豐富的語意描述，諸如文字轉寫、語者、情感、說話時場景等資訊。因此，文件摘要研究在最近十年來已將觸角延伸至語音文件摘要，常使用的研究題材包括了廣播新聞(Broadcast News)[19, 33-34]、語音郵件(Voice Mails)[35-36]、會議錄音(Meeting Recordings) [28, 31, 37]、演講紀錄(Lecture Recordings)[38]等。早期較直覺式的語音文件摘要作法是直接將語音文件利用自動語音辨識技術，將語音訊號轉換為文字轉寫，然後採取與文字文件摘要相同的技術以得到語音文件的自動摘要。然而，此舉卻突顯出在從事語音文件摘要時會面臨傳統文字文件摘要所未曾遭遇問題，例如語句中往往會帶有語音辨識錯誤、遲疑、重覆或無意義的語助詞等錯誤或不必要資訊。再者，因為語音文件的文字轉寫通常是以純文字(Plain Text)的方式呈現，缺乏像章節、標點符號、字體大小、句子邊界等結構資訊，以致於傳統文字文件摘要方法並無法正確地摘要出語音文件中重要文句的問題[2]。此外，語音文



















