

使用查詢意向探索與類神經網路於語音文件檢索之研究

Exploring Query Intent and Neural Network modeling Techniques for

Spoken Document Retrieval

羅天宏 Tien-Hong Lo, 陳映文 Ying-Wen Chen, 陳柏琳 Berlin Chen

國立臺灣師範大學資訊工程學系

Department of Computer Science and Information Engineering,

National Taiwan Normal University

E-mail: {teinhonglo, cliffchen, berlin}@ntnu.edu.tw

陳冠宇 Kuan-Yu Chen

國立台灣科技大學資訊工程系

Department of Computer Science and Information Engineering,

National Taiwan University of Science and Technology

kychen@mail.ntust.edu.tw

王新民 Hsin-Min Wang

中央研究院資訊科學研究所

Institute of Information Science,

Academia Sinica

whm@iis.sinica.edu.tw

摘要

伴隨著網際網路快速發展與多媒體資訊的大量增長，影音的傳遞與瀏覽越來越多並且成為我們日常生活的重要活動，這使得關於語音文件檢索(Spoken Document Retrieval, SDR)的研究成為一個有魅力的研究主題[1][2][3][4]。一般而言，SDR 的研究主要可分成兩大研究方向：第一個研究方向為建立具強健性的索引(Robust Indexing)以表達語音文件中詞彙和語意內涵，並且減緩語音辨識錯誤所造成的影響；第二個研究方向為發展有效的檢索模型(Effective Retrieval Models)，基於索引所代表的詞彙和語意內涵來量化使用者輸入的查詢(Query)和語音文件的相似程度，以協助使用者找到相關資訊，可分

為向量式模型[5]、統計式語言模型[6][7][8]。近年來，隨著類神經網路的突破，這類的方法也被大量應用在檢索的任務上[9][10]。雖然這幾年取得長足的進步，但如同傳統文字文件檢索，語音文件檢索也面臨輸入查詢過於簡短不能完整地表達使用者資訊需求的情況。因此，不少研究是利用使用者一次檢索後，將點擊的文件視為相關文章並做為查詢意向(Query Intent)[11]的依據，並在下一次提供更為精確的檢索結果，可視為與準相關回饋(Pseudo Relevance Feedback) [12]相似的流程，但必須進行兩次檢索的程序而會有耗時的問題。

基於上述的觀察，為了有效的查詢表示和即時性，我們發想了一個基於查詢意向應用在語音文件檢索的新穎方法，這樣的方法有兩個部份的貢獻。第一個部分，根據不同的查詢意向自動分辨查詢的種類，並使用相關性感知查詢模型(Neural Relevance-Aware Query model, NRM)推論出更有區別力(Discriminant)的查詢表示，豐富原先語意表達不足的查詢。不僅如此，相較於準相關回饋，不用進行兩次檢索，因此也附帶即時回饋的好處。第二個貢獻是在 NRM 架構下，分析查詢分群多寡對於系統效能的影響，我們認為不同主題的查詢應該各自分群訓練，強化模型的鑑別能力。實驗結果顯示，基於查詢效果下，我們的方法可進一步地提升訓練的有效性。

關鍵詞：查詢意向、類神經網路、準相關回饋

參考文獻

- [1] C. Chelba, T. J. Hazen, and M. Saraclar, “Retrieval and browsing of spoken content,” *IEEE Signal Processing Magazine*, 25(3), pp. 39–49, 2008.
- [2] L. S. Lee and B. Chen, “Spoken document understanding and organization,” *IEEE Signal Processing Magazine*, 22(5), pp. 42–60, 2005.
- [3] C. L. Huang, B. Ma, H. Li, and C.-H. Wu, “Speech indexing using semantic context inference,” in *Proceedings of INTERSPEECH*, pp. 717–720, 2011.
- [4] B. Chen, K.-Y. Chen, P.-N. Chen, and Y.-W. Chen, “Spoken document retrieval with unsupervised query modeling techniques,” *IEEE Transactions on Audio, Speech, and Language Processing* 20(9), pp. 2602–2612, 2012.
- [5] G. Salton, A. Wong, and C. S. Yang, “A vector space model for automatic indexing,” *Communications of the ACM*, 18(11), pp. 613–620, 1975

- [6] J. M. Ponte and W. B. Croft, “A language modeling approach to information retrieval,” in *Proceedings of SIGIR*, pp. 275–281, 1998.
- [7] F. Song and W. B. Croft, “A general language model for information retrieval,” in *Proceedings of CIKM*, pp. 316–321, 1999.
- [8] W. B. Croft and J. Lafferty (eds.), “Language modeling for information retrieval,” *Kluwer International Series on Information Retrieval*, 13, Kluwer Academic Publishers, 2003.
- [9] J.-F. Guo, Y. Fan, Q. Ai, and W. B. Croft, “A deep relevance matching model for ad-hoc retrieval,” in *Proceedings of CIKM*, pp. 55–64, 2016.
- [10] B. Mitra, F. Diaz, and N. Craswell, “Learning to match using local and distributed representations of text for web search,” in *Proceedings of WWW*, pp. 1291–1299, 2017.
- [11] Shen, D., Pan, R., Sun, J. T., Pan, J. J., Wu, K., Yin, J., & Yang, Q. (2006). Query enrichment for web-query classification. *ACM Transactions on Information Systems (TOIS)*, 24(3), 320-352.
- [12] C. Zhai and J. Lafferty, “Model-based feedback in the language modeling approach to information retrieval,” in *Proceedings of CIKM*, pp. 403–410, 2001.