

Journée d'étude sur les résumés multimédia
Paris, 17 Mars 2011
ATALA, AFCP, GDR-ISIS

Compte-Rendu

Cette réunion a attiré une quarantaine de participants issus des 3 communautés langage naturel, texte, et signal.

9h30 Conférence invitée: Horacio Saggion (Universitat Pompeu Fabra)
Can summaries help natural language processing? Study of the effect of text summarization in opinion classification and other tasks

En introduction, les méthodes utilisées en résumé automatique et les paradigmes d'évaluation sont passées en revue. Deux tâches courantes en traitement automatique des langues qui peuvent exploiter le résumé automatique sont ensuite exposées.

La première tâche est la coréférence entre plusieurs documents. Il s'agit de savoir si deux entités nommées (comme les personnes) citées dans plusieurs documents font référence à la même entité, sachant que les noms propres sont très ambigus. La méthode utilisée est de faire du clustering sur un ensemble de documents qui contiennent le même nom de personne jusqu'à ce qu'un seuil défini soit atteint. Les données pour l'évaluation proviennent du 1st Web People Search Evaluation) et la mesure appliquée est le F-score. L'utilisation de différents vecteurs de représentation du document (mots, entités, types d'entités, différents types de résumés,...) a été évaluée avec différents taux de compression, menant à un total d'environ 900 expériences. Généralement l'utilisation de résumés ne dégrade pas les résultats, mais n'améliore pas les résultats non plus.

La seconde tâche est la classification des opinions. Le lexique utilisé est SentiWordNet, en combinaison avec des SVM dans GATE. L'expérience consiste à comparer la catégorisation des revues (de banque, de film) et de leurs résumés. Pour les revues de banques, 160 conditions ont été testées avec plus de 7000 résumés. Beaucoup de conditions ont donné des résultats pires en utilisant les résumés que les documents complets, alors que certains en étaient proches. Dans les résultats préliminaires sur les critiques de films, il existe des catégorisations de sentiments qui sont meilleures sur les résumés.

10h30 Frederik Cailliau (Sinequa)
Le projet RPM2: résumé pluri-média, multi-documents et multi-opinions

Le résumé automatique est la question centrale du projet RPM2 (janvier 2008 – juin 2011). Le défi relevé par les partenaires du projet est de construire les méthodes et les outils pour faire du résumé sur le texte, l'audio et le vidéo. L'application visée est la structuration et la présentation de différentes sources d'information sous forme de résumé pour un agrégateur d'information (comme Wikio) ou un moteur de recherche (comme Sinequa).

Pour le résumé texte, un classement des phrases est faite avec des mesures de similarité, exploitant au passage une compression des phrases. Le choix des segments peut être guidé par l'opinion (positif – négatif) exprimé. Les résultats ont été évalués avec la mesure Rouge, et une méthode d'évaluation sans référence est en cours de développement. Deux corpus sont disponibles pour la communauté sur le site web. Le résumé audio est fait sur du vidéo afin d'être plus attractif et peut être orienté différemment selon les besoins. Deux méthodes ont été mises en œuvre (classique et zapping) et ont été évaluées. Le résumé vidéo se fait par la sélection des segments importants et la présentation d'affichettes. Les méthodes Video-MMR et AV-MMR ont été implémentées et évaluées avec la nouvelle mesure Vert.

11h30 Karaman Svebor (LaBRI)

Résumé en activités du quotidien de vidéos portées

Dans le cadre de la recherche sur la maladie d'Alzheimer, le projet IMMED propose un dispositif innovant permettant l'observation des activités du quotidien au domicile de patients pour l'évaluation de pertes d'autonomies. Une caméra fixée sur un gilet porté par le patient enregistre une vidéo en plan séquence d'un ensemble d'activités. La production d'un résumé automatique en termes d'activités permet au médecin d'observer efficacement les éléments d'intérêts.

12h Francois Portet (LIG)

Génération automatique de résumés textuels de données multimodales pour la prise de décision en unité néonatal

Les travaux présentés cadrent dans le projet anglais The BabyTalk Project (<http://www.csd.abdn.ac.uk/research/babytalk/>). Lors de la surveillance des patients, des données graphiques sont produites en grandes quantités. Pour faciliter l'interprétation de ces graphiques un résumé texte est produit qui les traduit en texte. Le défi est donc de générer un résumé de texte à partir de données graphiques. Afin de consolider la terminologie utilisée par les différents modules d'analyse des graphiques et de génération de texte, une ontologie a été créée. Elle oblige à fixer les concepts et permet de raisonner. Le résumé comporte environ 20 événements qui sont sélectionnés par un coefficient d'importance. Celui-ci est attribué selon des informations d'expert et des heuristiques. La génération du texte est faite avec SimpleNLG. L'évaluation a été faite sur place avec du personnel médical comme utilisateurs.

12h30 Emilie Guimier De Neef (Orange Labs)

Un outil de géolocalisation et de résumé automatique pour faciliter l'accès à l'information dans des corpus d'actualité

L'outil de résumé automatique multi-documents développé par Orange Labs enrichit les services proposés sur 2424actu, un service d'Orange. Ce site d'actualités est un agrégateur de contenus produits par des médias professionnels (50 partenaires). L'affichage cartographique est brièvement introduit : il exploite les entités nommées et la base Geonames pour projeter des clusters de documents sur une carte. Le résumé des documents a été développé dans le cadre d'un contrat Cifre (Aurélien Bossard). L'objectif est d'extraire les phrases qui maximisent la centralité (pertinence) et la diversité des

informations tout en minimisant la redondance. La méthode a été appelé CBSEAS et repose sur un calcul de similarité entre phrases et une classification des phrases. Par classe, une phrase représentative est extraite. On note la participation aux campagnes TAC 2008 et 2009. La classification non supervisée des phrases a été implémentée avec 2 mesures de centralité, comparée avec d'autres méthodes et évaluée avec la mesure Rouge sur le corpus RPM2.

14h15 Benoit Favre (LIF)

Evaluation du résumé automatique

Comme la traduction automatique, le problème de résumé automatique n'admet pas une solution unique. En effet, étant donné un ou plusieurs documents, chacun créera un résumé différent. Cette problématique est vraie que le médium à résumer ou le médium dans lequel le résumé est produit soit écrit, audio, ou vidéo. Dans ma présentation, je reviendrai sur les différentes méthodes d'évaluation de systèmes de résumé automatique qu'elles soient manuelles ou automatiques, intrinsèques ou applicatives, traitant du fond ou de la forme. Je dresserai un bilan de leurs forces et leurs faiblesses et lancerai la discussion sur ce thème essentiel pour mesurer l'avancement du résumé automatique multimédia.

14h45 Guillaume Gravier (IRISA)

Vers le résumé automatique de l'oral sans transcription

De récents travaux sur la découvertes de motifs récurrents dans les contenus sonores permettent de découvrir et de collecter des occurrences de mots réitérés dans les contenus parlé, sans recourir à des connaissances acoustique et linguistique a priori. Les motifs découverts dans les contenus parlés sont d'une part des événements liés à la production de la parole (inspirations, euh, etc.) et, d'autre part, des mots ou locutions. Ces derniers motifs peuvent servir de base à la construction d'un résumé sonore par sélection de motifs clés dans un premier temps puis de segments contenant ces motifs. La sélection de motifs clés peut se faire à l'aide de critères de fréquence ou d'un critère prosodique. Nous présentons une synthèse de récents travaux dans cette direction. Nous abordons tout d'abord la découverte de motifs sonores. Nous discutons ensuite leur sélection par des critères prosodiques. Enfin, nous illustrons le concept par une démonstration.

15h15 Geoffroy Peeters (IRCAM)

Génération automatique de résumés audio musicaux

Plusieurs techniques existent pour la génération automatique d'un résumé audio d'un morceau de musique. Dans tout les cas, l'objectif est d'obtenir un extrait audio court résumant le ou les instants clefs d'un morceau. La technique développée par l'Ircam repose sur l'analyse de la structure temporelle d'un morceau de musique en terme de répétition de segments (ou de séquences) au cours du temps. Celle-ci est ensuite utilisée pour générer le résumé.

Dans un premier temps, des descripteurs sont d'abord extraits du signal audio afin de représenter son contenu : MFCC, Chroma/PCP, SFM/SCM. Dans l'approche par état, cette suite temporelle des descripteurs est analysée afin d'obtenir une représentation sous forme de segments. Pour cela, des approches de type clustering (k-means ou hierarchical agglomerative clustering) ou HMM peuvent être

utilisée. Dans l'approche par séquence, une matrice de similarité est calculée. De ses diagonales (indiquant des répétitions) sont extraites les différentes séquences.

Une fois la structure estimée, les segments ou séquences clefs sont déterminés (fonction de leur nombre d'occurrence, de leur position temporelle, de leur sonie moyenne). Les trois segments les plus représentatifs sont finalement concaténés par une technique d'over-lap synchrone à la position de premiers temps. Pour cela un algorithme estime parallèlement la grille temporelle du morceau. Une évaluation de cette technique sur une base de 150 titres, montre que dans 95% des cas, un des trois extraits contient le titre du morceau considéré comme l'instant clef.

16h15 Abderrafih Lehman (Mining Essential)

Le résumé automatique de texte est-il maintenant mature pour une industrialisation logicielle ? L'expérience de la société Mining Essential

Abderrafih Lehman a présenté le logiciel de résumé automatique Essential Summarizer commercialisé par Mining Essential. Ce produit multi-plateforme est disponible en 20 langues et couvre 66 domaines. Il utilise les éléments discursifs du texte pour pondérer les phrases et générer un résumé automatique dont la longueur peut être modifiée par l'utilisateur en temps réel. Parmi les dernières nouveautés se trouvent la mise en évidence des phrases les plus importantes du texte par une police de caractères agrandie, l'intégration d'un traducteur automatique du résumé, le résumé de pages web et l'affichage de synonymes. Le logiciel peut être testé en ligne sur essential-mining.com.

16h45 Alina Elma Abduraman (Orange Labs)

Structuration intra-programmes de contenu TV

An important feature for automatic segmentation and structuring of TV programs is to provide tools that allow users to efficiently navigate within TV programs, that is, directly seek a certain part of a program or particular moments of interest. To achieve that, the original structure of each TV program needs to be recovered. In other terms, the start and end times of each part of the program have to be detected. This processing step cannot be performed manually when dealing with the huge amount of available TV programs.

This work addresses the unsupervised structuring of recurrent TV programs. These are programs that are periodically broadcasted, e.g. entertainment, shows, magazines... The proposed solution considers for each TV program a set of episodes. The idea is to automatically detect separators. These are short sequences that are inserted between different parts of the program, and that are generally repeated over the episodes or within the same episode. To achieve that, a repeated video sequence detection method is applied over the set of episodes. The detected repeated sequences are further analyzed in order to filter out those that are not separators. The effectiveness of the proposed solution has been shown on 3 different French TV programs.

17h15 Adriana Garboan (Telecom SudParis)

Video fingerprinting - a tool for multimedia automated summarization systems

Video fingerprinting can be best defined in relation with human fingerprinting: the human fingerprint can be seen as a human summary (a signature) that is unique for every person; similarly the video

fingerprint can be seen as a short video feature which can uniquely identify the videos. Although generally not connected to automated multimedia summarization, video fingerprinting can serve applications related to the same field. The challenge our research takes is video identification and retrieval in large databases.

In our study we developed an original fingerprinting method which is able to identify video queries of 20s long in a database. The principle is to exploit the mathematical modeling of the DWT coefficients in order to extract unique identifiers and to optimize the detection rule. The experiments proved the viability of the method: the probabilities of false alarm and of missed detection are both lower than 10^{-4} . The corpus consists of 3 hours of original video content, belonging to 7 different movies and combining indoor and outdoor scenes, unstable and arbitrary lighting conditions, still and high motion scenes. These 3h were further processed by some content preserving transformations, like conversion to gray scale, contrast decrease, sharpening, small rotations (2°), linear filtering (Gaussian filter) and geometric (StirMark random bending). Hence, the corpus sums up to 21 hours of video content.

17h45 **Clôture**