http://dx.doi.org/10.7236/JIWIT.2012.12.6.135
JIWIT 2012-6-17

사람 재식별: 학제간 연구 과제

People Re-identification: A Multidisciplinary Challenge

정동선*

Dong-Seon Cheng

요 약 인터넷의 확산과 정보 교환, 배포와 수집 기술에 대한 의존도의 증대로 과거와는 비교할 수 없는 대용량의 데이터가 생성되었다. 대용량 데이터를 식별하고 가려내는 작업은 가까운 미래에 오늘날의 컴퓨터 과학의 상당 부분을 새롭게 정의할 것으로 예상된다. 여러 관련 분야에서 반복되는 중요한 과제는 재식별의 문제이다. 광범위한 정의에서, 재식별 문제는 과거에 인식된 객체를 다시 식별하는 문제이다. 예를 들면, 여러 장소에 설치된 감시 카메라에 포착된 어떤 사람을 추적하는 문제가 이에 해당한다. 본 논문에서는 서로 다른 분야에서 이 과제를 어떻게 정의하고, 이 과제를 어떻게 해결하는가에 대해 비교 분석한다. 비디오 감시에서 사람 재식별, 텍스트 샘플에서 저자 식별, 사진 선호도에 따른 사용자 식별 등이 이에 포함된다. 본 논문은 또한 학제간 해결 방안이 장점을 지니는 상황에 대한 비전을 제시한다.

Abstract The wide diffusion of internet and the overall increased reliance on technology for information communication, dissemination and gathering have created an unparalleled mass of data. Sifting through this data is defining and will define in the foreseeable future a big part of contemporary computer science. Within this data, a growing proportion is given by personal information, which represents a unique opportunity to study human activities extensively and live. One important recurring challenge in many disciplines is the problem of people re-identification. In its broadest definition, re-identification is the problem of newly recognizing previously identified people, such as following an unknown person while he walks through many different surveillance cameras in different locations. Our goals is to review how several diverse disciplines define and meet this challenge, from person re-identification in video-surveillance to authorship attribution in text samples to distinguishing users based on their preferences of pictures. We further envision a situation where multidisciplinary solutions might be beneficial.

Key Words : people re-identification, video-surveillance, authorship attribution, computational aesthetics.

Ⅰ. Introduction

World-wide, our interconnected technological society is generating a huge amount of data. Due to cheap storage and communication means, an increasing proportion of this data is being kept indefinitely, be it for safety (surveillance), for further analysis (data mining for market research), or for other reasons (backup, convenience).

Through consumer-oriented internet based products,

*정회원, 한국외국어대학교 컴퓨터공학과
접수일자: 2012년 10월 18일, 수정일자: 2012년 11월 28일
게재확정일자 : 2012년 12월 14일

Received: 18 October 2012 / Revised: 28 November 2012 / Accepted: 14 December 2012
*Corresponding Author: cheng.dong.seon@gmail.com
Dept. of Computer Science & Engineering, HUFS, Korea

like messaging systems and social networks, users have lately introduced, mostly willingly, a large amount of personal information online, logging extensively – and often live – many diverse activities that characterize human behavior.

This represents a unique opportunity for disciplines that wish to analyze human activities. One recurring challenge for these disciplines is the fact that often the data can be anonymous (or anonymized intentionally), unformatted, or loosely tagged. The problem can be compounded by (partially) missing, erroneous or maliciously provided (with the intent of deceiving) data. Assuming supervised processing of the data is impractical or too expensive, a solution can be provided by re-identification.

People re-identification aims at labeling patterns of interest in a consistent way such that the same patterns, encountered independently in different settings, bear the same labels. For example, in video-surveillance, person re-identification[1] matches people's images in different locations over various non-overlapping camera views.

In this paper, we extend the concept of person re-identification from video-surveillance (Section II) to other research areas that analyze human activities, such as authorship attribution in textual chat conversations (Section III) and recognizing personal tastes in pictures selections (Section IV). Our goal is to show that the challenge of people re-identification can be interpreted more broadly but approached through similar pattern recognition tools.

In the conclusion, we further envision a situation where multi-modal multidisciplinary solutions might provide improved insights.

II. Person Re-identification

Person re-identification (re-id) in video surveillance commonly assume that the individuals observed through the cameras do not change their clothing within the observation period, and that finer biometric cues (face, fingerprint, gait) are unavailable[1].

Re-id is an important problem: it has been the focus of intense research in the last years, due to the distribution of challenging datasets, but its roots lie farther back, in object model design for human tracking. It is also pervasive, extending from the original video-surveillance field to the most recent photo-tagging domain.

The approach by Cheng et al.[2] takes inspiration from how humans perform re-id by selectively focusing on the body parts, looking for part-to-part correspondences. This is achieved through using Pictorial Structures (PS). PS essentially rely on two components: one capturing the local appearance of body parts, and the other representing an articulated body structure[3]. Inference in a PS involves finding the MAP spatial configuration of the parts, that is, the body pose (see Fig. 1).

그림 1. 그림구조로 구한 몸의 포즈 분석 예제.
Fig 1. Examples of body pose estimations obtained with Pictorial Structures.

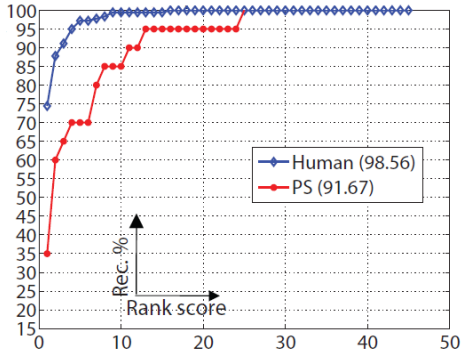

그림 2. 사람과 그림구조 접근법 사이의 CMC 그래프 비교
Fig 2. Comparison CMC graphs between humans and the Pictorial Structures approach[2].

With the body pose found, it is possible to calculate locally consistent appearance based descriptors that form a person's *signature*. This signature can be then searched and matched against eventho we do not give a name to the person. In video-surveillance, it is then possible to follow people from camera to camera by automatically searching for their signatures.

Re-id systems are usually evaluated by calculating a performance graph known as the cumulative matching characteristic (CMC), which represents the expectation of finding the correct match in the top n matches (Fig. 2) when sorting in decreasing order of similarity. It is clear that humans largely outperform the state-of-the-art algorithms, but, conversely, these algorithms can process more data cheaply.

The general framework of person re-identification can be applied mostly intact to other disciplines where the data is copious but anonymous, passively recorded but not volatile.

Ⅲ. Authorship Attribution

Authorship Attribution (AA) aims at automatically recognizing the author of a given text sample, based on the analysis of stylometric cues that can be split into five major groups: lexical, syntactic, structural, content-specific and idiosyncratic[4]. One important AA challenge is the identification of people in chat (or chat-like) conversations[5]. The wide diffusion of social media have created a situation where people interact through online identities using nick-names. By design or by users desire, these can hide their real identities and introduce the potential for interactions with malicious fake profiles.

In this context, re-identification can be seen as the problem of recognizing users by the way they write and interact: Cristani et al.[6] take into account the conversational nature of chat interactions, by defining new features based on turn-taking, probably the most salient aspect of spoken conversations that applies to chat interactions as well (see Table 2).

표 1. Cristani이 사용한 Stylometric 기능
Table 1. Stylometric features used by Cristani et al.[6] In bold, the conversational features.

No.	Feature	Range
1	# words	[0,260]
2	# emoticons	[0,40]
3	# emoticons per word	[0,1]
4	# emoticons per characters	[0,0.5]
5	# exclamation marks	[0,12]
6	# question marks	[0,406]
7	# characters	[0,1318]
8	average word length	[0,20]
9	# three points	[0,34]
10	# uppercase letters	[0,94]
11	# uppercase letters/#words	[0,290]
12	**turn duration**	[0,1800(sec.)]
13	**# return chars**	[1,20]
14	**# chars per second**	[0,20(ch./sec.)]
15	**# words per second**	[0,260]
16	**mimicry degree**	[0,1115]

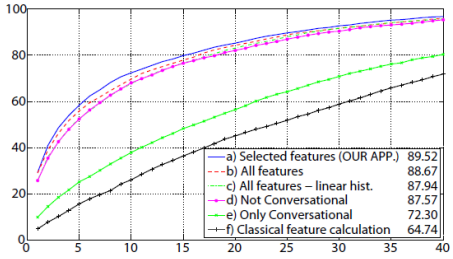

그림 3. 다른 기능 풀 사이의 CMC 그래프 비교
Fig 3. Comparison CMC graphs (and nAUC values) between different pools of features[6].

Re-id experiments were carried out on real-life dyadic Skype conversations, extracting users signatures from 30 "turns" of chat, where each turn is an interval of time during which only one user writes. By selecting a certain set of features, it is possible to obtain a very promising re-id performance (see Fig. 3). Each CMC graph can be summarized by a single number - the *normalized Area Under Curve* (nAUC in short), which measures overall accuracy.

From these and other similar results, we can see that it is possible to recognize us from the way we write and how we interact in a written conversation. This is important, because, after speech, writing is our most common expressive medium.

Ⅳ. Discriminating Visual Preferences

It is arguably accurate to say that we can be identified by the choices we make, especially if they are repeated and consistent. For example, people often get enjoyment from observing images and express preferences for some pictures over others. Lovato et al.[7] collect images from Flickr, a popular website where every user can select his/her preferred photos, by tagging them as "favorites". They then analyze the favorites sets of 200 users to infer about their personal aesthetics traits, that is, the features that distinguish the choices of different persons. These features represent personal taste - a user's signature, allowing to re-id the user from a set of 20 previously unseen images of an unknown user (see Fig. 4).

그림 4. 알려진 사용자와 정의되지 않은 이미지 set 사이의 매칭 예제

Fig 4. Example of matching a known user (left) and an unidentified set of images (right).

Quantitatively, this approach is able to reach a nAUC of 0.910. Qualitatively, the system is able to capture some of the coherence behind people's preferences, recognizing esthetic qualities that may guide their choices.

Ⅴ. Conclusion

Human activities are endlessly fascinating for what they reveal about our individuality and our societies. Not two people are ever alike, and this fact is exploited by biometric studies and applications to create and recognize identities. In many real world situations, though, the number of "known identities" is far outnumbered by the unknown ones - persons that have been detected but they are not actively sought or even named. Think about all the people catched in a network of video-surveillance cameras: the majority is of no interest, until something happens, like an emergency or a suspicious activity, and then the data is searched back for traces of events that lead to the crisis. A system that anticipates such needs could use an automatic re-identification algorithm to keep track of unknown identities passively. Other disciplines, like author attribution and image preferences, tell us that we can build identities based on what we write and how we write it and what we choose. This could be used to find people using different nick-names, or recommend choices or products based on our behaviors and preferences. There are also situations that cross disciplinary boundaries, and future applications might need multi-disciplinary capabilities. For example, video-chats would require a combination of video and conversational behavior recognition, with speech recognition as well.

In conclusion, re-identification provides a useful framework for the analysis of unlabeled mass data containing personal information of unknown or anonymized origin.

References

[1] G. Doretto, T. Sebastian, P. Tu, J. Rittscher. "Appearance-based person reidentification in camera networks: problem overview and current approaches." Journal of Ambient Intelligence and Humanized Computing, Vol.2, pp.127-151, 2011.

[2] D.S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, V. Murino. "Custom Pictorial Structures for Re-identication." Proc. of BMVC, 2011.

[3] P. F. Felzenszwalb and D. P. Huttenlocher. "Pictorial Structures for Object Recognition." Intl. J. on Computer Vision, 61(1):55 - 79, 2005.

[4] A. Abbasi and H. Chen. “Writeprints: A stylometric approach to identity-level identication and similarity detection in cyberspace.″ ACM Trans. on Information Systems, 26(2):1-29, 2008.

[5] A. Orebaugh and J. Allnutt. “Classification of instant messaging communications for forensics analysis.″ Social Networks, pp.22-28, 2009.

[6] M. Cristani, G. Roffo, C. Segalin, L. Bazzani, A. Vinciarelli, V. Murino. “Conversationally-inspired stylometric features for authorship attribution in instant messaging.″ In ACM Multimedia, 2012.

[7] P. Lovato, A. Perina, O. Zandona, A. Montagnini, N. Sebe, M. Bicego, M. Cristani, “Tell me what you like and I′ll tell you what you are: discriminating visual preferences by crowdsourcing.″ Proc. of the 11th Asian Conf. on Computer Vision, 2012.

저자 소개

Cheng Dong Seon(정회원)

- 2003년 : Laurea degree in Computer Science, University of Verona, Italy.
- 2008년 : Ph.D. degree in Computer Science, University of Verona, Italy.
- 2008년 ~ 2009년 : Post-doc researcher, EU-FP7 SIMBAD Project, Dept. of Computer Science, University of Verona, Italy.
- 2010년~2011년 : Post-doc researcher, EU-FP7 SAMURAI Project, Dept. of Computer Science, University of Verona, Italy.
- 2012년~Now : Assistant Professor, Dept. of Computer Science & Engineering, HUFS.

<Main Interests : Pattern Recognition, Computer Vision, Machine Learning, Image Processing, Video Surveillance>

※ This work was supported by Hankuk University of Foreign Studies Research Fund of 2012