Image processing for java / dot net

Check following Projects ,also check if any spelling mistakes before showing to your Guide:

Finger image quality assessment features – definitions and evaluation

Abstract: Finger image quality assessment is a crucial part of any system where a high biometric performance and user satisfaction is desired. Several algorithms measuring selected aspects of finger image quality have been proposed in the literature, yet only few of them have found their way into quality assessment algorithms used in practice. The authors provide comprehensive algorithm descriptions and make available implementations of adaptations of ten quality assessment algorithms from the literature which operates at the local or the global image level. They evaluate the performance on four datasets in terms of the capability in determining samples causing false non-matches and by their Spearman correlation with sample utility. The authors' evaluation shows that both the capability in rejecting samples causing false non-matches and the correlation between features varies depending on the dataset.

Extra Security Using Graphical Password to Data

Abstract—Many security primitives are based on hard mathematical problems. Using hard AI problems for security is emerging as an exciting new paradigm, but has been underexplored. In this paper, we present a new security primitive based on hard AI problems, namely, a novel family of graphical password systems built on top of Captcha technology, which we call Captcha as graphical passwords (CaRP). CaRP is both a Captcha and a graphical password scheme. CaRP addresses a number of security problems altogether, such as online guessing attacks, relay attacks, and, if combined with dual-view technologies, shoulder-surfing attacks. Notably, a CaRP password can be found only probabilistically by automatic online guessing attacks even if the password is in the search set. CaRP also offers a novel approach to address the well-known image hotspot problem in popular graphical password systems, such as PassPoints, that often leads to weak password choices. CaRP is not a panacea, but it offers reasonable security and usability and appears to fit well with some practical applications for improving online security.

An Efficient Privacy-Preserving RankedKeyword Search Method

Abstract—Cloud data owners prefer to outsource documents in an encrypted form for the purpose of privacy preserving. Therefore it is essential to develop efficient and reliable ciphertext search techniques. One challenge is that the relationship between documents will be normally concealed in the process of encryption, which will lead to significant search accuracy performance degradation. Also the volume of data in data centers has experienced a dramatic growth. This will make it even more challenging to design ciphertext search schemes that can provide efficient and reliable online information retrieval on large volume of encrypted data. In this paper, a hierarchical clustering method is proposed to support more search semantics and also to meet the demand for fast ciphertext search within a big data environment. The proposed hierarchical approach clusters the documents based on the minimum relevance threshold, and then partitions the resulting clusters into sub-clusters until the constraint on the maximum size of cluster is reached. In the search phase, this approach can reach a linear computational complexity against an exponential size increase of document collection. In order to verify the authenticity of search results, a structure called minimum hash sub-tree is designed in this paper. Experiments have been conducted using the collection set built from the IEEE Xplore. The results show that with a sharp increase of documents in the dataset the search time of the proposed method increases linearly whereas the search time of the traditional method increases exponentially. Furthermore, the proposed method has an advantage over the traditional method in the rank privacy and relevance of retrieved documents.

An Efficient SVD-Based Method for Image Denoising.

Abstract—Nonlocal self-similarity of images has attracted considerable interest in the field of image processing and has led to several state-of-the-art image denoising algorithms, such as block matching and 3-D, principal component analysis with local pixel grouping, patch-based locally optimal wiener, and spatially adaptive iterative singular-value thresholding. In this paper, we propose a computationally simple denoising algorithm using the nonlocal self-similarity and the low-rank approximation (LRA). The proposed method consists of three basic steps. First, our method classifies similar image patches by the block-matching technique to form the similar patch groups, which results in the similar patch groups to be low rank. Next, each group of similar patches is factorized by singular value decomposition (SVD) and estimated by taking only a few largest singular values and corresponding singular vectors. Finally, an initial denoised image is generated by aggregating all processed patches. For low-rank matrices, SVD can provide the optimal energy compaction in the least square sense. The proposed method exploits the optimal energy compaction property of SVD to lead an LRA of similar patch groups. Unlike other SVDbased methods, the LRA in SVD domain avoids learning the local basis for representing image patches, which usually is computationally expensive. The experimental results demonstrate that the proposed method can effectively reduce

noise and be competitive with the current state-of-the-art denoising algorithms in terms of both quantitative metrics and subjective visual quality.

Comparison of steganography using differents Algorithm.

Abstract : Steganography is an art to hide the existence of important information in a cover file. It is an information hiding technique which is used for sending and receiving confidential data over internet. Steganography is done in two part first is to embed data in regular computer file and the second part to extract that information. Secret data can be embed in various regular computer file but video files plays an important role by providing more embedding space. This paper will provide a survey of various research papers on video stegnography.

Object detection in video/ image.

Abstract : Object tracking is an important task within the field of computer vision. The proliferation of high-powered computers, the availability of high quality and inexpensive video cameras, and the interesting need for automated video analysis has generated a great deal of interest in object tracking.In its simplest form, tracking can be defined as a method of following an object through successive image frames to determine its relative movement with respect to other objects. In other words, a tracker assigns consistent labels to the tracked objects in different frames of video. One can simplify tracking by imposing constraints on the motion or appearance of objects. One can further constrain the object motion to be of constant velocity or acceleration based on prior information. Prior knowledge about the number and the size of objects, or the object appearance and shape can also be used to simplify the problem.

Handwritten Chinese Text Recognition by Integrating Multiple Contexts

Abstract —We describe Google's online handwriting recognition system that currently supports 22 scripts and 97 languages. The system's focus is on fast, high-accuracy text entry for mobile, touch-enabled devices. We use a combination of state-of-the-art components and combine them with novel additions in a flexible framework. This architecture allows us to easily transfer improvements between languages and scripts. This made it possible to build recognizers for languages that, to the best of our knowledge, are not handled by any other online handwriting recognition system. The approach also enabled us to use the same architecture both on very powerful machines for recognition in the cloud as well as on mobile devices with more limited computational power by

changing some of the settings of the system. In this paper we give a general overview of the system architecture and the novel components, such as unified time- and position-based input interpretation, trainable segmentation, minimum-error rate training for feature combination, and a cascade of pruning strategies. We present experimental results for different setups. The system is currently publicly available in several Google products, for example in Google Translate and as an input method for Android devices.

Learn to Personalized Image Search from the Photo Sharing Websites

Abstract—: As the amount of Web information grows rapidly, search engines must be able to retrieve information according to the user's preference. In this paper, we propose a new web search personalization approach that captures the user's interests and preferences in the form of concepts by mining search results and their clickthroughs. Due to the important role location information plays in mobile search, we separate concepts into content concepts and location concepts, and organize them into ontologies to create an ontology-based, multi-facet (OMF) prole to precisely capture the user's content and location interests and hence improve the search accuracy. Moreover, recognizing the fact that different users and queries may have different emphases on content and location information, we introduce the notion of content and location entropies to measure the amount of content and location information associated with a query, and click content and location entropies to measure how much the user is interested in the content and location information in the results. Accordingly, we propose to dene personalization effectiveness based on the entropies and use it to balance the weights between the content and location facets. Finally, based on the derived ontologies and personalization effectiveness, we train an SVM to adapt a personalized ranking function for re-ranking of future search. We conduct extensive experiments to compare the precision produced by our OMF proles and that of a baseline method. Experimental results show that OMF improves the precision signicantly compared to the baseline.

Enhance Security for online database.

Abstract — In this era due to unbelievable development in internet, various online attacks has been increased. From all such attacks most popular attack is phishing. This attacks are done for extracting confidential information such as banking information, passwords from unsuspecting victims for fraud purposes. Confidential data can’t be directly uploaded on website since it is risky. Here in this paper data is encrypted in video and visual cryptography for login purpose in our online database system for providing more security .

Face detection in video/image.

Abstract —This paper proposes a generic methodology for the semi-automatic generation of reliable position annotations for evaluating multi-camera people-trackers on large video data sets. Most of the annotation data are automatically computed, by estimating a consensus tracking result from multiple existing trackers and people detectors and classifying it as either reliable or not. A small subset of the data, composed of tracks with insufficient reliability, is verified by a human using a simple binary decision task, a process faster than marking the correct person position. The proposed framework is generic and can handle additional trackers. We present results on a data set of ~6 h captured by 4 cameras, featuring a person in a holiday flat, performing activities such as walking, cooking, eating, cleaning, and watching TV. When aiming for a tracking accuracy of 60 cm, 80% of all video frames are automatically annotated. The annotations for the remaining 20% of the frames were added after human verification of an automatically selected subset of data. This involved ~2.4 h of manual labor. According to a subsequent comprehensive visual inspection to judge the annotation procedure, we found 99% of the automatically annotated frames to be correct. We provide guidelines on how to apply the proposed methodology to new data sets. We also provide an exploratory study for the multi-target case, applied on the existing and new benchmark video sequences.

Elliptic Curve Cryptography.

Abstract — Secure and efficient data storage is needed in the cloud environment in modern era of information technology industry. In the present scenario the cloud verifies the authenticity of the cloud services without the knowledge of user's identity. The cloud provides massive data access directly through the internet. Centralized storage mechanism is followed here for effective accessing of data. Cloud service providers are normally acquires the software and hardware resources and the cloud consumers are avail the services through the internet access in lease basis. Cloud security is enhanced through cryptography technique applied to the cloud security to avoid vulnerability. The intractable computability is achieved in the cloud by using the public key cryptosystem. This paper proposed the approach of applying Hyper elliptic curve cryptography for data protection in the cloud with the small key size. The proposed system has the further advantage of eliminating intruder in cloud computing. Efficacy of the system is to provide the high security of the cloud data.

Vehicle detection in Aerial Surveillance.

Abstract

We present an automatic vehicle detection system for aerial surveillance in this paper. In this system, we escape from the stereotype and existing frameworks of vehicle detection in aerial surveillance, which are either region based or sliding window based. We design a pixel wise classification method for vehicle detection. The novelty lies in the fact that, in spite of performing pixel wise classification, relations among neighboring pixels in a region are preserved in the feature extraction process. We consider features including vehicle colors and local features. For vehicle color extraction, we utilize a color transform to separate vehicle colors and non-vehicle colors effectively. For edge detection, we apply moment preserving to adjust the thresholds of the Canny edge detector automatically, which increases the adaptability and the accuracy for detection in various aerial images. Afterward, a dynamic Bayesian network (DBN) is constructed for the classification purpose. We convert regional local features into quantitative observations that can be referenced when applying pixel wise classification via DBN. Experiments were conducted on a wide variety of aerial videos. The results demonstrate flexibility and good generalization abilities of the proposed method on a challenging data set with aerial surveillance images taken at different heights and under different camera angles.

Photo Morphing Detection.

ABSTRACT : In this digital world we come across many image processing software that produce doctored Images with high sophistication, which are manipulated in such a way that the tampering is not easily visible to naked eye. The authenticity of a digital image has become a challenging task due to the various tools present in the photo editing software packages. There are number of ways of tampering an Image, such as splicing two different images together, removal of objects from the image, addition of objects in the image, change of appearance of objects in the image or resizing the image. This Image Morphing detection technique detects traces of digital tampering in the complete absence of any form of digital watermark or signature and is therefore referred as passive. So there is a need for developing techniques to distinguish the original images from the manipulated ones, the genuine ones from the doctored ones. In this paper we describe a novel approach for detecting Image morphing. The new scheme is designed to detect any changes to a signal. We recognize that images from digital cameras contain traces of re-sampling as a result of using a color filter array with demosaicing algorithms. Our results show that the proposed scheme has a good accuracy in locating tampered pixels

Video watermarking by DCT algorithm.

ABSTRACT: Video data hiding is still an important research topic due to the design complexities involved. We propose a new video data hiding method that makes use of erasure correction capability of Repeat Accumulate codes and superiority of Forbidden Zone Data Hiding. Selective embedding is utilized in the proposed method to determine host signal samples suitable for data hiding. This method also contains a temporal synchronization scheme in order to withstand frame drop and insert attacks. The proposed framework is tested by typical broadcast material against MPEG-2, H.264 compression, frame-rate conversion attacks, as well as other well-known video data hiding methods. The decoding error values are reported for typical system parameters. The simulation results indicate that the framework can be successfully utilized in video data hiding applications.

Content Based Image Retrieval in Global server.

Abstract : The content-based image retrieval (CBIR) is the most acceptable and often used image retrieval method, because it can be used to manage image database efficiently and effectively. The CBIR methods usually retrieve the images by image features. In this paper, we exploit a region called affine invariant region (AIR) as an image feature to help effectively retrieving the images which have been attacked or processed. Moreover, we use vector quantization to reduce the features comparison for improving the retrieval efficiency. The experimental results show that the method with high recall and precision is promising.

Reversible Data Hiding: Advances in the Past Two Decades.

Abstract —In the past two decades, reversible data hiding (RDH), also referred to as lossless or invertible data hiding, has gradually become a very active research area in the field of data hiding. This has been verified by more and more papers on increasingly wide-spread subjects in the field of RDH research that have been published these days. In this survey paper the various RDH algorithms and researches have been classified into the following six categories: 1) RDH into image spatial domain, 2) RDH into image compressed domain (e.g., JPEG), 3) RDH suitable for image semi-fragile authentication, 4) RDH with image contrast enhancement, 5) RDH into encrypted images, which is expected to have wide application in the cloud computation, and 6) RDH into video and into audio. For each of these six categories, the history of technical developments, the current state of the arts, and the possible future researches are presented and discussed. It is expected that the RDH technology and its applications in the real word will continue to move ahead.

On the Properties of Non-media Digital Watermarking: A Review of State of the Art Techniques.

Abstract —Over the last 25 years, there has been much work on multimedia digital watermarking. In this domain, the primary limitation to watermark strength has been in its visibility. For multimedia watermarks, invisibility is defined in human terms (that is, in terms of human sensory limitations). In this paper, we review recent developments in the non-media applications of data watermarking, which have emerged over the last decade as an exciting new sub-domain. Since by definition, the intended receiver should be able to detect the watermark, we have to redefine invisibility in an acceptable way that is often application-specific and thus cannot be easily generalized. In particular, this is true when the data is not intended to be directly consumed by humans. For example, a loose definition of robustness might be in terms of the resilience of a watermark against normal host data operations, and of invisibility as resilience of the data interpretation against change introduced by the watermark. In our paper, we classify the data in terms of data mining rules on complex types of data such as time-series, symbolic sequences, data streams and so forth. We emphasize the challenges involved in non-media watermarking in terms of common watermarking properties including invisibility, capacity, robustness, and security. With the aid of a few examples of watermarking applications, we demonstrate these distinctions and we look at the latest research in this regard to make our argument clear and more meaningful. As the last aim, we look at the new challenges of digital watermarking that have arisen with the evolution of big data.

Remote Authentication via Biometrics: A Robust Video-Object Steganographic Mechanism Over Wireless Networks.

ABSTRACT In wireless communications, sensitive information is frequentlyexchanged, requiring remote authentication. Remote authentication involves the submission of encrypted information, along with visual and audio cues (facial images/videos, human voice, and so on). Nevertheless, Trojan horse and other attacks can cause serious problems, especially in the cases of remote examinations (in remote studying) or interviewing (for personnel hiring). This paper proposes a robust authentication mechanism based on semantic segmentation, chaotic encryption, and data hiding. Assuming that user X wants to be remotely authenticated, initially X's video object (VO) is automatically segmented, using a head-and-body detector. Next, one of X's biometric signals is encrypted by a chaotic cipher. Afterwards, the encrypted signal is inserted to the most signicant wavelet coefcients of the VO, using its qualied signicant wavelet trees (QSWTs). QSWTs provide both invisibility and signicant resistance against lossy transmission and compression, conditions that

are typical of wireless networks. Finally, the inverse discrete wavelet transform is applied to provide the stego-object. Experimental results regarding: 1) security merits of the proposed encryption scheme; 2) robustness to steganalytic attacks, to various transmission losses and JPEG compression ratios; and 3) bandwidth efciency measures indicate the promising performance of the proposed biometrics-based authentication scheme.

High-speed visual tracking with mixed rotation invariant description.

Abstract :

Visual target tracking is widely applied in visual surveillance, human–computer interaction, visual navigation and activity analysis. However, the response speed of conventional tracking systems is limited to <60 fps due to serial processing. Some researchers adopt parallel single-instruction multiple-data (SIMD) processors to speed up tracking algorithms [1–3]. However, these processors can only carry out simple algorithms such as background subtraction, segmentation and motion detection, thus they can only be applied to certain sceneries with a clean background. The local binary pattern (LBP) histogram of gradient (HOG) feature description is widely used in target detection and tracking [4, 5]. However, both the HOG and LBP histograms are rotation variant, which results in target shifting and tracking failure. In this Letter, we propose a mixed rotation invariant description (MRID)-based tracking algorithm and a novel high-speed visual tracking system. This MRID is invariant to rotation and illumination changes so that it achieves more robust tracking than previously reported fast tracking algorithms. The proposed tracking system integrates processors with pixel and row-level parallelism to speed up the tracking algorithm. The system with hierarchical parallelism can achieve over 1000 fps processing speed.

Multimodal BCIs: Target Detection, Multidimensional Control, and Awareness Evaluation in Patients With Disorder of Consciousness Despite rapid advances in the study of brain-computer.

ABSTRACT | Despite rapid advances in the study of brain– computer interfaces (BCIs) in recent decades, two fundamental challenges, namely, improvement of target detection performance and multidimensional control, continue to be major barriers for further development and applications. In this paper, we review the recent progress in multimodal BCIs (also called hybrid BCIs), which may provide potential solutions for addressing these challenges. In particular, improved target detection can be achieved by developing multimodal BCIs that utilize multiple brain patterns, multimodal

signals, or multisensory stimuli. Furthermore, multidimensional object control can be accomplished by generating multiple control signals from different brain patterns or signal modalities. Here, we highlight several representative multimodal BCI systems by analyzing their paradigm designs, detection/control methods, and experimental results. To demonstrate their practicality, we report several initial clinical applications of these multimodal BCI systems, including awareness evaluation/detection in patients with disorder of consciousness (DOC). As an evolving research area, the study of multimodal BCIs is increasingly requiring more synergetic efforts from multiple disciplines for the exploration of the underlying brainmechanisms, the design of new effective paradigms and means of neurofeedback, and the expansion of the clinical applications of these systems.

Ontology-Based Semantic Image Segmentation Using Mixture Models and Multiple CRFs.

Abstract —Semantic image segmentation is a fundamental yet challenging problem, which can be viewed as an extension of the conventional object detection with close relation to image segmentation and classification. It aims to partition images into non-overlapping regions that are assigned predefined semantic labels. Most of the existing approaches utilize and integrate lowlevel local features and high-level contextual cues, which are fed into an inference framework such as, the conditional random field (CRF). However, the lack of meaning in the primitives (i.e., pixels or superpixels) and the cues provides low discriminatory capabilities, since they are rarely object-consistent. Moreover, blind combinations of heterogeneous features and contextual cues exploitation through limited neighborhood relations in the CRFs tend to degrade the labeling performance. This paper proposes an ontology-based semantic image segmentation (OBSIS) approach that jointly models image segmentation and object detection. In particular, a Dirichlet process mixture model transforms the low-level visual space into an intermediate semantic space, which drastically reduces the feature dimensionality. These features are then individually weighed and independently learned within the context, using multiple CRFs. The segmentation of images into object parts is hence reduced to a classification task, where object inference is passed to an ontology model. This model resembles the way by which humans understand the images through the combination of different cues, context models, and rule-based learning of the ontologies. Experimental evaluations using the MSRC-21 and PASCAL VOC'2010 data sets show promising results.

Review of Video and Image Defogging Algorithms and Related Studies on Image Restoration and Enhancement.

ABSTRACT : Video and images acquired by a visual system are seriously degraded under hazy and foggy weather, which will affect the detection, tracking, and recognition of targets. Thus, restoring the true scene from such a foggy video or image is of signicance. The main goal of this paper was to summarize current video and image defogging algorithms. We rst presented a review of the detection and classication method of a foggy image. Then, we summarized existing image defogging algorithms, including image restoration algorithms, image contrast enhancement algorithms, and fusion-based defogging algorithms. We also presented current video defogging algorithms. We summarized objective image quality assessment methods that have been widely used for the comparison of different defogging algorithms, followed by an experimental comparison of various classical image defogging algorithms. Finally, we presented the problems of video and image defogging which need to be further studied.

Learning Sampling Distributions for Efficient Object Detection

Abstract —Object detection is an important task in computer vision and machine intelligence systems. Multistage particle windows (MPW), proposed by Gualdi et al., is an algorithm of fast and accurate object detection. By sampling particle windows (PWs) from a proposal distribution (PD), MPW avoids exhaustively scanning the image. Despite its success, it is unknown how to determine the number of stages and the number of PWs in each stage. Moreover, it has to generate too many PWs in the initialization step and it unnecessarily regenerates too many PWs around object-like regions. In this paper, we attempt to solve the problems of MPW. An important fact we used is that there is a large probability for a randomly generated PW not to contain the object because the object is a sparse event relative to the huge number of candidate windows. Therefore, we design a PD so as to efficiently reject the huge number of nonobject windows. Specifically, we propose the concepts of rejection, acceptance, and ambiguity windows and regions. Then, the concepts are used to form and update a dented uniform distribution and a dented Gaussian distribution. This contrasts to MPW which utilizes only on region of support. The PD of MPW is acceptance-oriented whereas the PD of our method (called iPW) is rejection-oriented. Experimental results on human and face detection demonstrate the efficiency and the effectiveness of the iPW algorithm. The source code is publicly accessible.

Captcha as Graphical Passwords—A New Security Primitive Based on Hard AI Problems.

Synopsis:

Many security primitives are based on hard mathematical problems. Using hard AI problems for security is emerging as an exciting new paradigm, but has been under-explored. In this paper, we present a new security primitive based on hard AI problems, namely, a novel family of graphical password systems built on top of Captcha technology, which we call Captcha as graphical passwords (CaRP). CaRP is both a Captcha and a graphical password scheme. CaRP addresses a number of security problems altogether, such as online guessing attacks, relay attacks, and, if combined with dual-view technologies, shoulder-surfing attacks. Notably, a CaRP password can be found only probabilistically by automatic online guessing attacks even if the password is in the search set. CaRP also offers a novel approach to address the well-known image hotspot problem in popular graphical password systems, such as PassPoints, that often leads to weak password choices. CaRP is not a panacea, but it offers reasonable security and usability and appears to fit well with some practical applications for improving online security.

An Adaptive Cloud Downloading Service.

Synopsis:

Video content downloading using the P2P approach is scalable, but does not always give good performance. Recently, subscription-based premium services have emerged, referred to as cloud downloading. In this service, the cloud storage and server caches user-interested content and updates the cache based on user downloading requests. If a requested video is not in the cache, the request is held in a waiting state until the cache is updated. We call this design server mode. An alternative design is to let the cloud server serve all downloading requests as soon as they arrive, behaving as a helper peer. We call this design helper mode. Our model and analysis show that both these designs are useful for certain operating regimes. The helper mode is good at handling a high request rate, while the server mode is good at scaling with video population size. We design an adaptive algorithm (AMS) to select the service mode automatically. Intuitively, AMS switches service

mode from server mode to helper mode when too many peers request blocked movies, and vice versa. The ability of AMS to achieve good performance in different operating regimes is validated by simulation .

Learn to Personalized Image Search from the Photo Sharing Websites.

Synopsis:

Increasingly developed social sharing websites like Flickr and Youtube allow users to create, share, annotate, and comment medias. The large-scale user-generated metadata not only facilitate users in sharing and organizing multimedia content, but provide useful information to improve media retrieval and management. Personalized search serves as one of such examples where the web search experience is improved by generating the returned list according to the modified user search intents. In this paper, we exploit the social annotations and propose a novel framework simultaneously considering the user and query relevance to learn to personalized image search. The basic premise is to embed the user preference and query-related search intent into user-specific topic spaces. Since the users' original annotation is too sparse for topic modeling, we need to enrich users' annotation pool before user-specific topic spaces construction. The proposed framework contains two components: (1) a ranking-based multicorrelation tensor factorization model is proposed to perform annotation prediction, which is considered as users' potential annotations for the images; (2) we introduce user-specific topic modeling to map the query relevance and user preference into the same user-specific topic space. For performance evaluation, two resources involved with users' social activities are employed. Experiments on a large-scale Flickr dataset demonstrate the effectiveness of the proposed method.

Learning and Recognition of On-Premise Signs from Weakly Labeled Street View Images.

Synopsis:

Camera-enabled mobile devices are commonly used as interaction platforms for linking the user's virtual and physical worlds in numerous research and commercial applications, such as serving an augmented reality interface for mobile information retrieval. The various application scenarios give rise to a key technique of daily life visual object recognition. On-premise signs (OPSs), a popular form of commercial advertising, are widely used in our living life. The OPSs often exhibit great visual diversity (e.g., appearing in arbitrary size),

accompanied with complex environmental conditions (e.g., foreground and background clutter). Observing that such real-world characteristics are lacking in most of the existing image data sets, in this paper, we first proposed an OPS data set, namely OPS-62, in which totally 4649 OPS images of 62 different businesses are collected from Google's Street View. Further, for addressing the problem of real-world OPS learning and recognition, we developed a probabilistic framework based on the distributional clustering, in which we proposed to exploit the distributional information of each visual feature (the distribution of its associated OPS labels) as a reliable selection criterion for building discriminative OPS models. Experiments on the OPS-62 data set demonstrated the outperformance of our approach over the state-of-the-art probabilistic latent semantic analysis models for more accurate recognitions and less false alarms, with a significant 151.28% relative improvement in the average recognition rate. Meanwhile, our approach is simple, linear, and can be executed in a parallel fashion, making it practical and scalable for large-scale multimedia applications..

An extended visual cryptography scheme without pixel expansion for halftone images.

Synopsis:

Visual cryptography is a secret sharing scheme which uses images distributed as shares such that, when the shares are superimposed, a hidden secret image is revealed. In extended visual cryptography, the share images are constructed to contain meaningful cover images, thereby providing opportunities for integrating visual cryptography and biometric security techniques. In this paper, we propose a method for processing halftone images that improves the quality of the share images and the recovered secret image in an extended visual cryptography scheme for which the size of the share images and the recovered image is the same as for the original halftone secret image. The resulting scheme maintains the perfect security of the original extended visual cryptography approach.

Trust modeling in social tagging of multimedia content.

Synopsis:

Tagging in online social networks is very popular these days, as it facilitates search and retrieval of multimedia content. However, noisy and spam annotations often make it difficult to perform an efficient search. Users may make mistakes in tagging and irrelevant tags and content may be maliciously added for advertisement or self-promotion. This article surveys recent advances in techniques for combatting such noise and spam in social tagging. We classify the state-of-the-art approaches into a few categories and study representative

examples in each. We also qualitatively compare and contrast them and outline open issues for future research.

Personalized Geo-Specific Tag Recommendation for Photos on Social Websites.

Synopsis:

Social tagging becomes increasingly important to organize and search large-scale community-contributed photos on social websites. To facilitate generating high-quality social tags, tag recommendation by automatically assigning relevant tags to photos draws particular research interest. In this paper, we focus on the personalized tag recommendation task and try to identify user-preferred, geo-location-specific as well as semantically relevant tags for a photo by leveraging rich contexts of the freely available community-contributed photos. For users and geo-locations, we assume they have different preferred tags assigned to a photo, and propose a subspace learning method to individually uncover the both types of preferences. The goal of our work is to learn a unified subspace shared by the visual and textual domains to make visual features and textual information of photos comparable. Considering the visual feature is a lower level representation on semantics than the textual information, we adopt a progressive learning strategy by additionally introducing an intermediate subspace for the visual domain, and expect it to have consistent local structure with the textual space. Accordingly, the unified subspace is mapped from the intermediate subspace and the textual space respectively. We formulate the above learning problems into a united form, and present an iterative optimization with its convergence proof. Given an untagged photo with its geo-location to a user, the user-preferred and the geo-location-specific tags are found by the nearest neighbor search in the corresponding unified spaces. Then we combine the obtained tags and the visual appearance of the photo to discover the semantically and visually related photos, among which the most frequent tags are used as the recommended tags. Experiments on a large-scale data set collected from Flickr verify the effectivity of the proposed solution.

Beyond Text QA: Multimedia Answer Generation by Harvesting Web Information.

Synopsis:

Community question answering (cQA) services have gained popularity over the past years. It not only allows community members to post and answer questions but also enables general users to seek information from a comprehensive set of well-answered questions.

However, existing cQA forums usually provide only textual answers, which are not informative enough for many questions. In this paper, we propose a scheme that is able to enrich textual answers in cQA with appropriate media data. Our scheme consists of three components: answer medium selection, query generation for multimedia search, and multimedia data selection and presentation. This approach automatically determines which type of media information should be added for a textual answer. It then automatically collects data from the web to enrich the answer. By processing a large set of QA pairs and adding them to a pool, our approach can enable a novel multimedia question answering (MMQA) approach as users can find multimedia answers by matching their questions with those in the pool. Different from a lot of MMQA research efforts that attempt to directly answer questions with image and video data, our approach is built based on community-contributed textual answers and thus it is able to deal with more complex questions. We have conducted extensive experiments on a multi-source QA dataset. The results demonstrate the effectiveness of our approach.

Bootstrapping Visual Categorization With Relevant Negatives.

Learning classifiers for many visual concepts are important for image categorization and retrieval. As a classifier tends to misclassify negative examples which are visually similar to positive ones, inclusion of such misclassified and thus relevant negatives should be stressed during learning. User-tagged images are abundant online, but which images are the relevant negatives remains unclear. Sampling negatives at random is the de facto standard in the literature. In this paper, we go beyond random sampling by proposing Negative Bootstrap. Given a visual concept and a few positive examples, the new algorithm iteratively finds relevant negatives. Per iteration, we learn from a small proportion of many user-tagged images, yielding an ensemble of meta classifiers. For efficient classification, we introduce Model Compression such that the classification time is independent of the ensemble size. Compared with the state of the art, we obtain relative gains of 14% and 18% on two present-day benchmarks in terms of mean average precision. For concept search in one million images, model compression reduces the search time from over 20 h to approximately 6 min. The effectiveness and efficiency, without the need of manually labeling any negatives, make negative bootstrap appealing for learning better visual concept classifiers.

Circular Reranking for Visual Search.

Synopsis:

Search reranking is regarded as a common way to boost retrieval precision. The problem nevertheless is not trivial especially when there are multiple features or modalities to be

considered for search, which often happens in image and video retrieval. This paper proposes a new reranking algorithm, named circular reranking, that reinforces the mutual exchange of information across multiple modalities for improving search performance, following the philosophy that strong performing modality could learn from weaker ones, while weak modality does benefit from interacting with stronger ones. Technically, circular reranking conducts multiple runs of random walks through exchanging the ranking scores among different features in a cyclic manner. Unlike the existing techniques, the reranking procedure encourages interaction among modalities to seek a consensus that are useful for reranking. In this paper, we study several properties of circular reranking, including how and which order of information propagation should be configured to fully exploit the potential of modalities for reranking. Encouraging results are reported for both image and video retrieval on Microsoft Research Asia Multimedia image dataset and TREC Video Retrieval Evaluation 2007-2008 datasets, respectively.

Hierarchical Super-Resolution-Based Inpainting.

Synopsis:

This paper introduces a novel framework for examplar-based inpainting. It consists in performing first the inpainting on a coarse version of the input image. A hierarchical super-resolution algorithm is then used to recover details on the missing areas. The advantage of this approach is that it is easier to inpaint low-resolution pictures than high-resolution ones. The gain is both in terms of computational complexity and visual quality. However, to be less sensitive to the parameter setting of the inpainting method, the low-resolution input picture is inpainted several times with different configurations. Results are efficiently combined with a loopy belief propagation and details are recovered by a single-image super-resolution algorithm. Experimental results in a context of image editing and texture synthesis demonstrate the effectiveness of the proposed method. Results are compared to five state-of-the-art inpainting methods.

Image Authentication Using Stochastic Diffusion.Synopsis:

This paper considers an approach to encrypted information hiding based on Stochastic Diffusion for encrypting digital images coupled with the application of a Least Significant Bit (LSB) method for information embedding. After providing a brief summary of various information hiding methods based on spatial and transform domain techniques, two new methods are introduced. The first of these considers a binary image watermarking algorithm for hiding an image in a single host image which is based on binarization of the encrypted data. The second method extends this approach to solving the problem of 24-bit image hiding in three host images which generates a near perfect reconstruction after decryption. Both methods make use of a 'hidden code' technique to randomize the order of the embedded bits and the location (in the image plane) of the LSBs which make the embedded

information more robust to attack. Details of the algorithms developed are provided and examples are given, which have application in the field of covert cryptography and the authentication of full colour images for copyright protection and Data Rights Management.

Scalable Face Image Retrieval using Attribute-Enhanced Sparse Codewords.

Synopsis:

Photos with people (e.g., family, friends, celebrities, etc.) are the major interest of users. Thus, with the exponentially growing photos, large-scale content-based face image retrieval is an enabling technology for many emerging applications. In this work, we aim to utilize automatically detected human attributes that contain semantic cues of the face photos to improve content-based face retrieval by constructing semantic codewords for efficient large-scale face retrieval. By leveraging human attributes in a scalable and systematic framework, we propose two orthogonal methods named attribute-enhanced sparse coding and attribute-embedded inverted indexing to improve the face retrieval in the offline and online stages. We investigate the effectiveness of different attributes and vital factors essential for face retrieval. Experimenting on two public datasets, the results show that the proposed methods can achieve up to 43.5% relative improvement in MAP compared to the existing methods..

Understanding the External Links of Video.

Synopsis:

Recently, many video sharing sites provide external links so that their video or audio contents can be embedded into external web sites. For example, users can copy the embedded URLs of the videos of YouTube and post the URL links on their own blogs. Clearly, the purpose of such function is to increase the distribution of the videos and the associated advertisement. Does this function fulfill its purpose and what is the quantification? In this paper, we provide a comprehensive measurement study and analysis on these external links to answer these two questions. With the traces collected from two major video sharing sites, YouTube and Youku of China, we show that the external links have various impacts on the popularity of the video sharing sites. More specifically, for videos that have been uploaded for eight months in Youku, around 15% of views can come from external links. Some contents are densely linked. For example, comedy videos can attract more than 800 external links on average. We also study the relationship between the external links and the internal links. We show that there are correlations; for example, if a video is popular itself, it is likely to have a large number of external links. Another observation we find is that the external links usually have a higher impact on Youku than

that of YouTube. We conjecture that it is more likely that the external links have higher impact for a regional site than a worldwide site.

A Compressive Sensing based Secure Watermark Detection and Privacy Preserving Storage.

Synopsis:

Privacy is a critical issue when the data owners outsource data storage or processing to a third party computing service, such as the cloud. In this paper, we identify a cloud computing application scenario that requires simultaneously performing secure watermark detection and privacy preserving multimedia data storage. We then propose a compressive sensing (CS)-based framework using secure multiparty computation (MPC) protocols to address such a requirement. In our framework, the multimedia data and secret watermark pattern are presented to the cloud for secure watermark detection in a CS domain to protect the privacy. During CS transformation, the privacy of the CS matrix and the watermark pattern is protected by the MPC protocols under the semi-honest security model. We derive the expected watermark detection performance in the CS domain, given the target image, watermark pattern, and the size of the CS matrix (but without the CS matrix itself). The correctness of the derived performance has been validated by our experiments. Our theoretical analysis and experimental results show that secure watermark detection in the CS domain is feasible. Our framework can also be extended to other collaborative secure signal processing and data-mining applications in the cloud.

LLSURE: Local Linear SURE-Based Edge-Preserving Image Filtering.

Synopsis:

In this paper, we propose a novel approach for performing high-quality edge-preserving image filtering. Based on a local linear model and using the principle of Stein's unbiased risk estimate as an estimator for the mean squared error from the noisy image only, we derive a simple explicit image filter which can filter out noise while preserving edges and fine-scale details. Moreover, this filter has a fast and exact linear-time algorithm whose computational complexity is independent of the filtering kernel size; thus, it can be applied to real time image processing tasks. The experimental results demonstrate the effectiveness of the new filter for various computer vision applications, including noise reduction, detail smoothing and enhancement, high dynamic range compression, and flash/no-flash denoising.

Altered Fingerprints: Analysis and Detection.

Synopsis:

The widespread deployment of Automated Fingerprint Identification Systems (AFIS) in law enforcement and border control applications has heightened the need for ensuring that these systems are not compromised. While several issues related to fingerprint system security have been investigated, including the use of fake fingerprints for masquerading identity, the problem of fingerprint alteration or obfuscation has received very little attention. Fingerprint obfuscation refers to the deliberate alteration of the fingerprint pattern by an individual for the purpose of masking his identity. Several cases of fingerprint obfuscation have been reported in the press. Fingerprint image quality assessment software (e.g., NFIQ) cannot always detect altered fingerprints since the implicit image quality due to alteration may not change significantly. The main contributions of this paper are: 1) compiling case studies of incidents where individuals were found to have altered their fingerprints for circumventing AFIS, 2) investigating the impact of fingerprint alteration on the accuracy of a commercial fingerprint matcher, 3) classifying the alterations into three major categories and suggesting possible countermeasures, 4) developing a technique to automatically detect altered fingerprints based on analyzing orientation field and minutiae distribution, and 5) evaluating the proposed technique and the NFIQ algorithm on a large database of altered fingerprints provided by a law enforcement agency. Experimental results show the feasibility of the proposed approach in detecting altered fingerprints and highlight the need to further pursue this problem.

A Novel Video Steganography based on Non-uniform Rectangular Partition.

Synopsis:

This paper proposes a novel Video Steganography which can hide an uncompressed secret video stream in a host video stream with almost the same size. Each frame of the secret video will be Non-uniform rectangular partitioned and the partitioned codes obtained can be an encrypted version of the original frame. These codes will be hidden in the Least 4 Significant Bits of each frames of the host video. Experimental results showed that this algorithm can hide a same-size video in the host video without obvious distortion in the host video.

A New Iterative TriclassThresholding Technique in Image Segmentation.

Synopsis:

We present a new method in image segmentation that is based on Otsu's method but iteratively searches for subregions of the image for segmentation, instead of treating the full image as a whole region for processing. The iterative method starts with Otsu's threshold and computes the mean values of the two classes as separated by the threshold. Based on the Otsu's threshold and the two mean values, the method separates the image into three classes instead of two as the standard Otsu's method does. The first two classes are determined as the foreground and background and they will not be processed further. The third class is denoted as a to-be-determined (TBD) region that is processed at next iteration. At the succeeding iteration, Otsu's method is applied on the TBD region to calculate a new threshold and two class means and the TBD region is again separated into three classes, namely, foreground, background, and a new TBD region, which by definition is smaller than the previous TBD regions. Then, the new TBD region is processed in the similar manner. The process stops when the Otsu's thresholds calculated between two iterations is less than a preset threshold. Then, all the intermediate foreground and background regions are, respectively, combined to create the final segmentation result. Tests on synthetic and real images showed that the new iterative method can achieve better performance than the standard Otsu's method in many challenging cases, such as identifying weak objects and revealing fine structures of complex objects while the added computational cost is minimal.

Local Directional Number Pattern for Face Analysis: Face and Expression Recognition.

Synopsis:

This paper proposes a novel local feature descriptor, local directional number pattern (LDN), for face analysis, i.e., face and expression recognition. LDN encodes the directional information of the face's textures (i.e., the texture's structure) in a compact way, producing a more discriminative code than current methods. We compute the structure of each micro-pattern with the aid of a compass mask that extracts directional information, and we encode such information using the prominent direction indices (directional numbers) and sign-which allows us to distinguish among similar structural patterns that have different intensity transitions. We divide the face into several regions, and extract the distribution of the LDN features from them. Then, we concatenate these features into a feature vector, and we use it as a face descriptor. We perform several experiments in which our descriptor performs consistently under illumination, noise, expression, and time lapse variations. Moreover, we test our descriptor with different masks to analyze its performance in different face analysis tasks.

AnnoSearch: Image Auto-Annotation by Search.

Synopsis:

Although it has been studied for several years by computer vision and machine learning communities, image annotation is still far from practical. In this paper, we present AnnoSearch, a novel way to annotate images using search and data mining technologies. Leveraging the Web-scale images, we solve this problem in two-steps: 1) searching for semantically and visually similar images on the Web, 2) and mining annotations from them. Firstly, at least one accurate keyword is required to enable text-based search for a set of semantically similar images. Then content-based search is performed on this set to retrieve visually similar images. At last, annotations are mined from the descriptions (titles, URLs and surrounding texts) of these images. It worth highlighting that to ensure the efficiency, high dimensional visual features are mapped to hash codes which significantly speed up the content-based search process. Our proposed approach enables annotating with unlimited vocabulary, which is impossible for all existing approaches. Experimental results on real web images show the effectiveness and efficiency of the proposed algorithm.

On the Typical Statistic Features for Image Blind Steganalysis.

Synopsis:

Multimedia content is a suitable carrier for secret communication. This paper focuses on the steganalysis technique which aims to get the forensic of secrecy existing in multimedia carriers. A key concern for designing a blind steganalysis algorithm is the selection of statistic features. The Probability Density Function (PDF) moment and Characteristic Function (CF) moment are two typical kinds of statistic features commonly used in blind steganalysis. And generally, the features are computed from the subbands of transform domains, such as the wavelet coefficient subbands, the prediction subbands of wavelet coefficients, the prediction error subbands of wavelet coefficients, the wavelet coefficient subbands of image noise, and the log prediction error subbands of wavelet coefficients. To decide which feature is more sensitive to message embedding and useful for steganalysis is important and urgent. Till now, few works have focused on this topic, and they can only give some experimental results without theoretical analysis. Additionally, few frequency subbands have been investigated. To solve this problem, this paper reviews existing feature computing algorithms, compares the two kinds of features, the PDF moments and the CF moments, by analyzing the change trends of the statistic distribution parameters of various frequency subbands before and after message embedding, and so that provides a theoretical basis for the steganalysis feature selection and extraction. These theoretical results are further confirmed by experimental results. This is the first work to provide thorough theoretical analysis on so many feature computing algorithms. It is expected to

provide valuable information to researchers or engineers working in the field of steganography forensics or steganalysis.

An Iranian License Plate Recognition System Based on Color Features.

Synopsis:

In this paper, an Iranian vehicle license plate recognition system based on a new localization approach, which is modified to reflect the local context, is proposed, along with a hybrid classifier that recognizes license plate characters. The method presented here is based on a modified template-matching technique by the analysis of target color pixels to detect the location of a vehicle's license plate. A modified strip search enables localization of the standard color-geometric template utilized in Iran and several European countries. This approach uses periodic strip search to find the hue of each pixel on demand. In addition, when a group of target pixels is detected, it is analyzed to verify that its shape and aspect ratio match those of the standard license plate. In addition to being scale and rotation invariant, this method avoids time-consuming image algorithms and transformations for the whole image pixels, such as resizing and Hough, Fourier, and wavelet transforms, thereby cutting down the detection response time. License plate characters are recognized by a hybrid classifier that comprises a decision tree and a support vector machine with a homogeneous fifth-degree polynomial kernel. The performance detection rate and the overall system performance achieved are 96% and 94%, respectively.

Perceptual Video Coding Based on SSIM-Inspired.

Synopsis:

We propose a perceptual video coding framework based on the divisive normalization scheme, which is found to be an effective approach to model the perceptual sensitivity of biological vision, but has not been fully exploited in the context of video coding. At the macroblock (MB) level, we derive the normalization factors based on the structural similarity (SSIM) index as an attempt to transform the discrete cosine transform domain frame residuals to a perceptually uniform space. We further develop an MB level perceptual mode selection scheme and a frame level global quantization matrix optimization method. Extensive simulations and subjective tests verify that, compared with the H.264/AVC video coding standard, the proposed method can achieve significant gain in terms of rate-SSIM performance and provide better visual quality.

Compressed-Sensing-Enabled Video Streaming for Wireless Multimedia Sensor Networks.

Synopsis:

This paper presents the design of a networked system for joint compression, rate control and error correction of video over resource-constrained embedded devices based on the theory of Compressed Sensing (CS). The objective of this work is to design a cross-layer system that jointly controls the video encoding rate, the transmission rate, and the channel coding rate to maximize the received video quality. First, compressed sensing-based video encoding for transmission over Wireless Multimedia Sensor Networks (WMSNs) is studied. It is shown that compressed sensing can overcome many of the current problems of video over WMSNs, primarily encoder complexity and low resiliency to channel errors. A rate controller is then developed with the objective of maintaining fairness among different videos while maximizing the received video quality. It is shown that the rate of Compressed Sensed Video (CSV) can be predictably controlled by varying only the compressed sensing sampling rate. It is then shown that the developed rate controller can be interpreted as the iterative solution to a convex optimization problem representing the optimization of the rate allocation across the network. The error resiliency properties of compressed sensed images and videos are then studied, and an optimal error detection and correction scheme is presented for video transmission over lossy channels. Finally, the entire system is evaluated through simulation and test bed evaluation. The rate controller is shown to outperform existing TCP-friendly rate control schemes in terms of both fairness and received video quality. The test bed results show that the rates converge to stable values in real channels.

Steganographic Embedding in JPEG Images with Visual Criterion.

Synopsis:

In this paper, we present a new information hiding scheme in JPEG images to achieve a good embedding efficiency considering visual criterion. We construct an embedding impact model based on human visual system, and then assign each cover element a flipping cost which would be the key parameter during the embedding procedure. In this way, the proposed method can minimize the total embedding impact via Viterbi algorithm, meanwhile improve the visual quality of the stego medium. The experimental results later show that the proposed information hiding system can perform well in different types of images.

As-Projective-As-Possible Image Stitching with Moving DLT.

Synopsis:

The success of commercial image stitching tools often leads to the impression that image stitching is a “solved problem”. The reality, however, is that many tools give unconvincing results when the input photos violate fairly restrictive imaging assumptions; the main two being that the photos correspond to views that differ purely by rotation, or that the imaged scene is effectively planar. Such assumptions underpin the usage of 2D projective transforms or homographies to align photos. In the hands of the casual user, such conditions are often violated, yielding misalignment artifacts or “ghosting” in the results. Accordingly, many existing image stitching tools depend critically on post-processing routines to conceal ghosting. In this paper, we propose a novel estimation technique called Moving Direct Linear Transformation (Moving DLT) that is able to tweak or fine-tune the projective warp to accommodate the deviations of the input data from the idealized conditions. This produces as-projective-as-possible image alignment that significantly reduces ghosting without compromising the geometric realism of perspective image stitching. Our technique thus lessens the dependency on potentially expensive postprocessing algorithms. In addition, we describe how multiple as-projective-as-possible warps can be simultaneously refined via bundle adjustment to accurately align multiple images for large panorama creation.

Selecting a Reference High Resolution for Fingerprint Recognition Using Minutiae and Pores.

Synopsis:

High-resolution automated fingerprint recognition systems (AFRSs) offer higher security because they are able to make use of level-3 features, such as pores, that are not available in lower resolution (<; 500-dpi) images. One of the main parameters affecting the quality of a digital fingerprint image and issues such as cost, interoperability, and performance of an AFRS is the choice of image resolution. In this paper, we identify the optimal resolution for an AFRS using the two most representative fingerprint features: minutiae and pores. We first designed a multiresolution fingerprint acquisition device to collect fingerprint images at multiple resolutions and captured fingerprints at various resolutions but at a fixed image size. We then carried out a theoretical analysis to identify the minimum required resolution for fingerprint recognition using minutiae and pores. After experiments on our collected fingerprint images and applying three requirements for the proportions of minutiae and pores that must be retained in a fingerprint image, we recommend a reference resolution of 800 dpi. Subsequent tests have further confirmed the proposed reference resolution.

DEFENSES AGAINST LARGE SCALE ONLINE PASSWORD GUESSING ATTACKS BY USING PERSUASIVE CLICK POINTS .

Synopsis:

Usable security has unique usability challenges because the need for security often means that standard human-computer-interaction approaches cannot be directly applied. An important usability goal for authentication systems is to support users in selecting better passwords. Users often create memorable passwords that are easy for attackers to guess, but strong system-assigned passwords are difficult for users to remember. So researchers of modern days have gone for alternative methods wherein graphical pictures are used as passwords. Graphical passwords essentially use images or representation of images as passwords. Human brain is good in remembering picture than textual character. There are various graphical password schemes or graphical password software in the market. However, very little research has been done to analyze graphical passwords that are still immature. There for, this project work merges persuasive cued click points and password guessing resistant protocol. The major goal of this work is to reduce the guessing attacks as well as encouraging users to select more random, and difficult passwords to guess. Well known security threats like brute force attacks and dictionary attacks can be successfully abolished using this method.

Virtualized Screen: A Third Element for CloudMobile Convergence.

Synopsis:

Mobile and cloud computing have emerged as the new computing platforms and are converging into a powerful cloud-mobile computing platform. This article envisions a virtualized screen as a new dimension in such a platform to further optimize the overall computing experience for users. In a virtualized screen, screen rendering is done in the cloud, and delivered as images to the client for interactive display. This enables thin-client mobile devices to enjoy many computationally intensive and graphically rich services. Technical challenges are discussed and addressed. Two novel cloud-mobile applications, Cloud Browser and Cloud Phone, are presented to demonstrate the advantages of such a virtualized screen.

Corruptive Artifacts Suppression for Example-Based ColorTransfer.

Synopsis:

Example-based color transfer is a critical operation in image editing but easily suffers from some corruptive artifacts in the mapping process. In this paper, we propose a novel unified color transfer framework with corruptive artifacts suppression, which performs iterative probabilistic color mapping with self-learning filtering scheme and multiscale detail manipulation scheme in minimizing the normalized Kullback-Leibler distance. First, an iterative probabilistic color mapping is applied to construct the mapping relationship between the reference and target images. Then, a self-learning filtering scheme is applied into the transfer process to prevent from artifacts and extract details. The transferred output and the extracted multi-levels details are integrated by the measurement minimization to yield the final result. Our framework achieves a sound grain suppression, color fidelity and detail appearance seamlessly. For demonstration, a series of objective and subjective measurements are used to evaluate the quality in color transfer. Finally, a few extended applications are implemented to show the applicability of this framework.

Visual Object Tracking Based on Local Steering Kernels and Color Histograms.

Synopsis:

In this paper, we propose a visual object tracking framework, which employs an appearance-based representation of the target object, based on local steering kernel descriptors and color histogram information. This framework takes as input the region of the target object in the previous video frame and a stored instance of the target object, and tries to localize the object in the current frame by finding the frame region that best resembles the input. As the object view changes over time, the object model is updated, hence incorporating these changes. Color histogram similarity between the detected object and the surrounding background is employed for background subtraction. Experiments are conducted to test the performance of the proposed framework under various conditions. The proposed tracking scheme is proven to be successful in tracking objects under scale and rotation variations and partial occlusion, as well as in tracking rather slowly deformable articulated objects.

Exploring the design space of social network-based Sybil defenses.

Synopsis:

Recently, there has been significant research interest in leveraging social networks to defend against Sybil attacks. While much of this work may appear similar at first glance, existing social network-based Sybil defense schemes can be divided into two categories: Sybil detection and Sybil tolerance. These two categories of systems both leverage global properties of the underlying social graph, but they rely on different assumptions and provide different guarantees: Sybil detection schemes are application-independent and rely only on the graph structure to identify Sybil identities, while Sybil tolerance schemes rely on application-specific information and leverage the graph structure and transaction history to bound the leverage an attacker can gain from using multiple identities. In this paper, we take a closer look at the design goals, models, assumptions, guarantees, and limitations of both categories of social network-based Sybil defense systems.

Localization of License Plate Number Using Dynamic Image Processing Techniques And Genetic Algorithms .

Synopsis:

In this research, the design of a new genetic algorithm (GA) is introduced to detect the locations of license plate (LP) symbols. An adaptive threshold method is applied to overcome the dynamic changes of illumination conditions when converting the image into binary. Connected component analysis technique (CCAT) is used to detect candidate objects inside the unknown image. A scale-invariant geometric relationship matrix is introduced to model the layout of symbols in any LP that simplifies system adaptability when applied in different countries. Moreover, two new crossover operators, based on sorting, are introduced, which greatly improve the convergence speed of the system. Most of the CCAT problems, such as touching or broken bodies, are minimized by modifying the GA to perform partial match until reaching an acceptable fitness value. The system is implemented using MATLAB and various image samples are experimented with to verify the distinction of the proposed system. Encouraging results with 98.4% overall accuracy are reported for two different datasets having variability in orientation, scaling, plate location, illumination, and complex background. Examples of distorted plate images are successfully detected due to the independency on the shape, color, or location of the plate.

FeatureMatch: A General ANNF Estimation Technique and its Applications.

Synopsis:

In this paper, we propose FeatureMatch, a generalised approximate nearest-neighbour field (ANNF) computation framework, between a source and target image. The proposed algorithm can estimate ANNF maps between any image pairs, not necessarily related. This generalisation is achieved through appropriate spatial-range transforms. To compute ANNF maps, global colour adaptation is applied as a range transform on the source image. Image patches from the pair of images are approximated using low-dimensional features, which are used along with KD-tree to estimate the ANNF map. This ANNF map is further improved based on image coherency and spatial transforms. The proposed generalisation, enables us to handle a wider range of vision applications, which have not been tackled using the ANNF framework. We illustrate two such applications namely: 1) optic disk detection and 2) super resolution. The first application deals with medical imaging, where we locate optic disks in retinal images using a healthy optic disk image as common target image. The second application deals with super resolution of synthetic images using a common source image as dictionary. We make use of ANNF mappings in both these applications and show experimentally that our proposed approaches are faster and accurate, compared with the state-of-the-art techniques.

Model-Based Edge Detector for Spectral Imagery Using Sparse Spatiospectral Masks.

Synopsis:

Two model-based algorithms for edge detection in spectral imagery are developed that specifically target capturing intrinsic features such as isoluminant edges that are characterized by a jump in color but not in intensity. Given prior knowledge of the classes of reflectance or emittance spectra associated with candidate objects in a scene, a small set of spectral-band ratios, which most profoundly identify the edge between each pair of materials, are selected to define a edge signature. The bands that form the edge signature are fed into a spatial mask, producing a sparse joint spatiospectral nonlinear operator. The first algorithm achieves edge detection for every material pair by matching the response of the operator at every pixel with the edge signature for the pair of materials. The second algorithm is a classifier-enhanced extension of the first algorithm that adaptively accentuates distinctive features before applying the spatiospectral operator. Both algorithms are extensively verified using spectral imagery from the airborne hyperspectral imager and from a dots-in-a-well midinfrared imager. In both cases, the multicolor gradient (MCG) and the hyperspectral/spatial detection of edges (HySPADE) edge detectors are used as a benchmark for comparison. The results demonstrate that the proposed algorithms outperform the MCG and HySPADE edge detectors in accuracy, especially when isoluminant edges are present. By requiring only a few bands as input to the spatiospectral operator, the algorithms enable significant levels of data compression in band selection. In the presented examples, the required operations per pixel are reduced by a factor of 71 with respect to those required by the MCG edge detector.

Improving Color Constancy by Photometric Edge Weighting.

Synopsis:

Edge-based color constancy methods make use of image derivatives to estimate the illuminant. However, different edge types exist in real-world images, such as material, shadow, and highlight edges. These different edge types may have a distinctive influence on the performance of the illuminant estimation. Therefore, in this paper, an extensive analysis is provided of different edge types on the performance of edge-based color constancy methods. First, an edge-based taxonomy is presented classifying edge types based on their photometric properties (e.g., material, shadow-geometry, and highlights). Then, a performance evaluation of edge-based color constancy is provided using these different edge types. From this performance evaluation, it is derived that specular and shadow edge types are more valuable than material edges for the estimation of the illuminant. To this end, the (iterative) weighted Gray-Edge algorithm is proposed in which these edge types are more emphasized for the estimation of the illuminant. Images that are recorded under controlled circumstances demonstrate that the proposed iterative weighted Gray-Edge algorithm based on highlights reduces the median angular error with approximately 25 percent. In an uncontrolled environment, improvements in angular error up to 11 percent are obtained with respect to regular edge-based color constancy.

Modeling of Speaking Rate Influences on Mandarin Speech Prosody and Its Application to Speaking Rate-controlled TTS.

Synopsis:

A new data-driven approach to building a speaking rate-dependent hierarchical prosodic model (SR-HPM), directly from a large prosody-unlabeled speech database containing utterances of various speaking rates, to describe the influences of speaking rate on Mandarin speech prosody is proposed. It is an extended version of the existing HPM model which contains 12 sub-models to describe various relationships of prosodic-acoustic features of speech signal, linguistic features of the associated text, and prosodic tags representing the prosodic structure of speech. Two main modifications are suggested. One is designing proper normalization functions from the statistics of the whole database to compensate the influences of speaking rate on all prosodic-acoustic features. Another is modifying the HPM training to let its parameters be speaking-rate dependent. Experimental results on a large Mandarin read speech corpus showed that the parameters of the SR-HPM together with these feature normalization functions interpreted the effects of speaking

rate on Mandarin speech prosody very well. An application of the SR-HPM to design and implement a speaking rate-controlled Mandarin TTS system is demonstrated. The system can generate natural synthetic speech for any given speaking rate in a wide range of 3.4-6.8 syllables/sec. Two subjective tests, MOS and preference test, were conducted to compare the proposed system with the popular HTS system. The MOS scores of the proposed system were in the range of 3.58-3.83 for eight different speaking rates, while they were in 3.09-3.43 for HTS. Besides, the proposed system had higher preference scores (49.8%-79.6%) than those (9.8%-30.7%) of HTS. This confirmed the effectiveness of the speaking rate control method of the proposed TTS system.

Fingerprint Compression Based on Sparse Representation.

Synopsis:

A new fingerprint compression algorithm based on sparse representation is introduced. Obtaining an overcomplete dictionary from a set of fingerprint patches allows us to represent them as a sparse linear combination of dictionary atoms. In the algorithm, we first construct a dictionary for predefined fingerprint image patches. For a new given fingerprint images, represent its patches according to the dictionary by computing l^0-minimization and then quantize and encode the representation. In this paper, we consider the effect of various factors on compression results. Three groups of fingerprint images are tested. The experiments demonstrate that our algorithm is efficient compared with several competing compression techniques (JPEG, JPEG 2000, and WSQ), especially at high compression ratios. The experiments also illustrate that the proposed algorithm is robust to extract minutiae.

Offline Text-Independent Writer Identification Based on Scale Invariant Feature Transform.

Synopsis:

This paper proposes a novel offline text-independent writer identification method based on scale invariant feature transform (SIFT), composed of training, enrollment, and identification stages. In all stages, an isotropic LoG filter is first used to segment the handwriting image into word regions (WRs). Then, the SIFT descriptors (SDs) of WRs and the corresponding scales and orientations (SOs) are extracted. In the training stage, an SD codebook is constructed by clustering the SDs of training samples. In the enrollment stage, the SDs of the input handwriting are adopted to form an SD signature (SDS) by looking up the SD codebook and the SOs are utilized to generate a scale and orientation histogram (SOH). In the identification stage, the SDS and SOH of the input handwriting are extracted and matched with the enrolled ones for identification. Experimental results on six public data

sets (including three English data sets, one Chinese data set, and two hybrid-language data sets) demonstrate that the proposed method outperforms the state-of-the-art algorithms.

IntentSearch:Capturing User Intention for One-Click Internet Image Search.

Synopsis:

Web-scale image search engines (e.g., Google image search, Bing image search) mostly rely on surrounding text features. It is difficult for them to interpret users' search intention only by query keywords and this leads to ambiguous and noisy search results which are far from satisfactory. It is important to use visual information in order to solve the ambiguity in text-based image retrieval. In this paper, we propose a novel Internet image search approach. It only requires the user to click on one query image with minimum effort and images from a pool retrieved by text-based search are reranked based on both visual and textual content. Our key contribution is to capture the users' search intention from this one-click query image in four steps. 1) The query image is categorized into one of the predefined adaptive weight categories which reflect users' search intention at a coarse level. Inside each category, a specific weight schema is used to combine visual features adaptive to this kind of image to better rerank the text-based search result. 2) Based on the visual content of the query image selected by the user and through image clustering, query keywords are expanded to capture user intention. 3) Expanded keywords are used to enlarge the image pool to contain more relevant images. 4) Expanded keywords are also used to expand the query image to multiple positive visual examples from which new query specific visual and textual similarity metrics are learned to further improve content-based image reranking. All these steps are automatic, without extra effort from the user. This is critically important for any commercial web-based image search engine, where the user interface has to be extremely simple. Besides this key contribution, a set of visual features which are both effective and efficient in Internet image search are designed. Experimental evaluation shows that our approach significantly improves the precision of top-ranked images and also the user experi- nce.

Photometric Stereo Using Sparse Bayesian Regression for General Diffuse Surfaces.

Synopsis:

Most conventional algorithms for non-Lambertian photometric stereo can be partitioned into two categories. The first category is built upon stable outlier rejection techniques while assuming a dense Lambertian structure for the inliers, and thus performance degrades when general diffuse regions are present. The second utilizes complex reflectance

representations and non-linear optimization over pixels to handle non-Lambertian surfaces, but does not explicitly account for shadows or other forms of corrupting outliers. In this paper, we present a purely pixel-wise photometric stereo method that stably and efficiently handles various non-Lambertian effects by assuming that appearances can be decomposed into a sparse, non-diffuse component (e.g., shadows, specularities, etc.) and a diffuse component represented by a monotonic function of the surface normal and lighting dot-product. This function is constructed using a piecewise linear approximation to the inverse diffuse model, leading to closed-form estimates of the surface normals and model parameters in the absence of non-diffuse corruptions. The latter are modeled as latent variables embedded within a hierarchical Bayesian model such that we may accurately compute the unknown surface normals while simultaneously separating diffuse from non-diffuse components. Extensive evaluations are performed that show state-of-the-art performance using both synthetic and real-world images.

How to Estimate the Regularization Parameter for Spectral Regression Discriminant Analysis and its Kernel Version?

Synopsis:

Spectral regression discriminant analysis (SRDA) has recently been proposed as an efficient solution to large-scale subspace learning problems. There is a tunable regularization parameter in SRDA, which is critical to algorithm performance. However, how to automatically set this parameter has not been well solved until now. So this regularization parameter was only set to be a constant in SRDA, which is obviously suboptimal. This paper proposes to automatically estimate the optimal regularization parameter of SRDA based on the perturbation linear discriminant analysis (PLDA). In addition, two parameter estimation methods for the kernel version of SRDA are also developed. One is derived from the method of optimal regularization parameter estimation for SRDA. The other is to utilize the kernel version of PLDA. Experiments on a number of publicly available databases demonstrate the effectiveness of the proposed methods for face recognition, spoken letter recognition, handwritten digit recognition, and text categorization.

Quality Assessment of Stereoscopic 3D Image Compression by Binocular Integration Behaviors.

Synopsis:

The objective approaches of 3D image quality assessment play a key role for the development of compression standards and various 3D multimedia applications. The quality assessment of 3D images faces more new challenges, such as asymmetric stereo compression, depth perception, and virtual view synthesis, than its 2D counterparts. In addition, the widely used 2D image quality metrics (e.g., PSNR and SSIM) cannot be directly applied to deal with these newly introduced challenges. This statement can be verified by the low correlation between the computed objective measures and the subjectively measured mean opinion scores (MOSs), when 3D images are the tested targets. In order to meet these newly introduced challenges, in this paper, besides traditional 2D image metrics, the binocular integration behaviors-the binocular combination and the binocular frequency integration, are utilized as the bases for measuring the quality of stereoscopic 3D images. The effectiveness of the proposed metrics is verified by conducting subjective evaluations on publicly available stereoscopic image databases. Experimental results show that significant consistency could be reached between the measured MOS and the proposed metrics, in which the correlation coefficient between them can go up to 0.88. Furthermore, we found that the proposed metrics can also address the quality assessment of the synthesized color-plus-depth 3D images well. Therefore, it is our belief that the binocular integration behaviors are important factors in the development of objective quality assessment for 3D images.

Robust Face-Name Graph Matching for Movie Character Identification.

Synopsis:

Automatic face identification of characters in movies has drawn significant research interests and led to many interesting applications. It is a challenging problem due to the huge variation in the appearance of each character. Although existing methods demonstrate promising results in clean environment, the performances are limited in complex movie scenes due to the noises generated during the face tracking and face clustering process. In this paper we present two schemes of global face-name matching based framework for robust character identification. The contributions of this work include the following. 1) A noise insensitive character relationship representation is incorporated. 2) We introduce an edit operation based graph matching algorithm. 3) Complex character changes are handled by simultaneously graph partition and graph matching. 4) Beyond existing character

identification approaches, we further perform an in-depth sensitivity analysis by introducing two types of simulated noises. The proposed schemes demonstrate state-of-the-art performance on movie character identification in various genres of movies.

Robust Semi-Automatic Depth Map Generation in Unconstrained Images and Video Sequences for 2D to Stereoscopic 3D Conversion.

Synopsis:

We describe a system for robustly estimating synthetic depth maps in unconstrained images and videos, for semi-automatic conversion into stereoscopic 3D. Currently, this process is automatic or done manually by rotoscopers. Automatic is the least labor intensive, but makes user intervention or error correction difficult. Manual is the most accurate, but time consuming and costly. Noting the merits of both, a semi-automatic method blends them together, allowing for faster and accurate conversion. This requires user-defined strokes on the image, or over several keyframes for video, corresponding to a rough estimate of the depths. After, the rest of the depths are determined, creating depth maps to generate stereoscopic 3D content, with Depth Image Based Rendering to generate the artificial views. Depth map estimation can be considered as a multi-label segmentation problem: each class is a depth. For video, we allow the user to label only the first frame, and we propagate the strokes using computer vision techniques. We combine the merits of two well-respected segmentation algorithms: Graph Cuts and Random Walks. The diffusion from Random Walks, with the edge preserving of Graph Cuts should give good results. We generate good quality content, more suitable for perception, compared to a similar framework

Image Classification Using Multiscale Information Fusion Based on Saliency Driven Nonlinear Diffusion Filtering.

Synopsis:

In this paper, we propose saliency driven image multiscale nonlinear diffusion filtering. The resulting scale space in general preserves or even enhances semantically important structures such as edges, lines, or flow-like structures in the foreground, and inhibits and smoothes clutter in the background. The image is classified using multiscale information fusion based on the original image, the image at the final scale at which the diffusion process converges, and the image at a midscale. Our algorithm emphasizes the foreground features, which are important for image classification. The background image regions, whether considered as contexts of the foreground or noise to the foreground, can be globally handled by fusing information from different scales. Experimental tests of the effectiveness of the multiscale space for the image classification are conducted on the following publicly available datasets: 1) the PASCAL 2005 dataset; 2) the Oxford 102 flowers dataset; and 3) the Oxford 17 flowers dataset, with high classification rates.

Sharing Visual Secrets in Single Image Random Dot Stereograms.

Synopsis:

Visual cryptography schemes (VCSs) generate random and meaningless shares to share and protect secret images. Conventional VCSs suffer from a transmission risk problem because the noise-like shares will raise the suspicion of attackers and the attackers might intercept the transmission. Previous research has involved in hiding shared content in halftone shares to reduce these risks, but this method exacerbates the pixel expansion problem and visual quality degradation problem for recovered images. In this paper, a binocular VCS (BVCS), called the ((2,n)) -BVCS, and an encryption algorithm are proposed to hide the shared pixels in the single image random dot stereograms (SIRDSs). Because the SIRDSs have the same 2D appearance as the conventional shares of a VCS, this paper tries to use SIRDSs as cover images of the shares of VCSs to reduce the transmission risk of the shares. The encryption algorithm alters the random dots in the SIRDSs according to the construction rule of the ((2,n)) -BVCS to produce nonpixel-expansion shares of the BVCS. Altering the dots in a SIRDS will degrade the visual quality of the reconstructed 3D objects. Hence, we propose an optimization model that is based on the visual quality requirement of SIRDSs to develop construction rules for a ((2,n)) -BVCS that maximize the contrast of the recovered image in the BVCS.

Robust Watermarking of Compressed and Encrypted JPEG2000 Images.

Synopsis:

Digital asset management systems (DAMS) generally handle media data in a compressed and encrypted form. It is sometimes necessary to watermark these compressed encrypted media items in the compressed-encrypted domain itself for tamper detection or ownership declaration or copyright management purposes. It is a challenge to watermark these compressed encrypted streams as the compression process would have packed the information of raw media into a low number of bits and encryption would have randomized the compressed bit stream. Attempting to watermark such a randomized bit stream can cause a dramatic degradation of the media quality. Thus it is necessary to choose an encryption scheme that is both secure and will allow watermarking in a predictable manner in the compressed encrypted domain. In this paper, we propose a robust watermarking algorithm to watermark JPEG2000 compressed and encrypted images. The encryption algorithm we propose to use is a stream cipher. While the proposed technique embeds watermark in the compressed-encrypted domain, the extraction of watermark can be done in the decrypted domain. We investigate in detail the embedding capacity, robustness, perceptual quality and security of the proposed algorithm, using these watermarking schemes: Spread Spectrum (SS), Scalar Costa Scheme Quantization Index Modulation (SCS-QIM), and Rational Dither Modulation (RDM).

Single-Image Superresolution of Natural Stochastic Textures Based on Fractional Brownian Motion.

Synopsis:

Texture enhancement presents an ongoing challenge, in spite of the considerable progress made in recent years. Whereas most of the effort has been devoted so far to enhancement of regular textures, stochastic textures that are encountered in most natural images, still pose an outstanding problem. The purpose of enhancement of stochastic textures is to recover details, which were lost during the acquisition of the image. In this paper, a texture model, based on fractional Brownian motion (fBm), is proposed. The model is global and does not entail using image patches. The fBm is a self-similar stochastic process. Self-similarity is known to characterize a large class of natural textures. The fBm-based model is evaluated and a single-image regularized superresolution algorithm is derived. The proposed algorithm is useful for enhancement of a wide range of textures. Its performance is compared with single-image superresolution methods and its advantages are highlighted.

LBP-Based Edge-Texture Features for Object Recognition.

Synopsis:

This paper proposes two sets of novel edge-texture features, Discriminative Robust Local Binary Pattern (DRLBP) and Ternary Pattern (DRLTP), for object recognition. By investigating the limitations of Local Binary Pattern (LBP), Local Ternary Pattern (LTP) and Robust LBP (RLBP), DRLBP and DRLTP are proposed as new features. They solve the problem of discrimination between a bright object against a dark background and vice-versa inherent in LBP and LTP. DRLBP also resolves the problem of RLBP whereby LBP codes and their complements in the same block are mapped to the same code. Furthermore, the proposed features retain contrast information necessary for proper representation of object contours that LBP, LTP, and RLBP discard. Our proposed features are tested on seven challenging data sets: INRIA Human, Caltech Pedestrian, UIUC Car, Caltech 101, Caltech 256, Brodatz, and KTH-TIPS2-a. Results demonstrate that the proposed features outperform the compared approaches on most data sets.

View-invariant action recognition based on Artificial Neural Networks .

Synopsis:

In this paper, a novel view invariant action recognition method based on neural network representation and recognition is proposed. The novel representation of action videos is based on learning spatially related human body posture prototypes using self organizing maps. Fuzzy distances from human body posture prototypes are used to produce a time invariant action representation. Multilayer perceptrons are used for action classification. The algorithm is trained using data from a multi-camera setup. An arbitrary number of cameras can be used in order to recognize actions using a Bayesian framework. The proposed method can also be applied to videos depicting interactions between humans, without any modification. The use of information captured from different viewing angles leads to high classification performance. The proposed method is the first one that has been tested in challenging experimental setups, a fact that denotes its effectiveness to deal with most of the open issues in action recognition.

Semisupervised Biased Maximum Margin Analysis for Interactive Image Retrieval.

Synopsis:

With many potential practical applications, content-based image retrieval (CBIR) has attracted substantial attention during the past few years. A variety of relevance feedback

(RF) schemes have been developed as a powerful tool to bridge the semantic gap between low-level visual features and high-level semantic concepts, and thus to improve the performance of CBIR systems. Among various RF approaches, support-vector-machine (SVM)-based RF is one of the most popular techniques in CBIR. Despite the success, directly using SVM as an RF scheme has two main drawbacks. First, it treats the positive and negative feedbacks equally, which is not appropriate since the two groups of training feedbacks have distinct properties. Second, most of the SVM-based RF techniques do not take into account the unlabeled samples, although they are very helpful in constructing a good classifier. To explore solutions to overcome these two drawbacks, in this paper, we propose a biased maximum margin analysis (BMMA) and a semisupervised BMMA (SemiBMMA) for integrating the distinct properties of feedbacks and utilizing the information of unlabeled samples for SVM-based RF schemes. The BMMA differentiates positive feedbacks from negative ones based on local analysis, whereas the SemiBMMA can effectively integrate information of unlabeled samples by introducing a Laplacianregularizer to the BMMA. We formally formulate this problem into a general subspace learning task and then propose an automatic approach of determining the dimensionality of the embedded subspace for RF. Extensive experiments on a large real-world image database demonstrate that the proposed scheme combined with the SVM RF can significantly improve the performance of CBIR systems.

User-aware Image Tag Refinement via Ternary Semantic Analysis.

Synopsis:

Large-scale user contributed images with tags are easily available on photo sharing websites. However, the noisy or incomplete correspondence between the images and tags prohibits them from being leveraged for precise image retrieval and effective management. To tackle the problem of tag refinement, we propose a method of Ranking based Multi-correlation Tensor Factorization (RMTF), to jointly model the ternary relations among user, image, and tag, and further to precisely reconstruct the user-aware image-tag associations as a result. Since the user interest or background can be explored to eliminate the ambiguity of image tags, the proposed RMTF is believed to be superior to the traditional solutions, which only focus on the binary image-tag relations. During the model estimation, we employ a ranking based optimization scheme to interpret the tagging data, in which the pair-wise qualitative difference between positive and negative examples is used, instead of the point-wise 0/1 confidence. Specifically, the positive examples are directly decided by the observed user-image-tag interrelations, while the negative ones are collected with respect to the most semantically and contextually irrelevant tags. Extensive experiments on a benchmark Flickr dataset demonstrate the effectiveness of the proposed solution for tag

refinement. We also show attractive performances on two potential applications as the by-products of the ternary relation analysis.

Learning Layouts for Single-Page Graphic Designs.

Synopsis:

This paper presents an approach for automatically creating graphic design layouts using a new energy-based model derived from design principles. The model includes several new algorithms for analyzing graphic designs, including the prediction of perceived importance, alignment detection, and hierarchical segmentation. Given the model, we use optimization to synthesize new layouts for a variety of single-page graphic designs. Model parameters are learned with Nonlinear Inverse Optimization (NIO) from a small number of example layouts. To demonstrate our approach, we show results for applications including generating design layouts in various styles, retargeting designs to new sizes, and improving existing designs. We also compare our automatic results with designs created using crowdsourcing and show that our approach performs slightly better than novice designers.

Separable Reversible Data Hiding in Encrypted Image.

Synopsis:

This work proposes a novel scheme for separable reversible data hiding in encrypted images. In the first phase, a content owner encrypts the original uncompressed image using an encryption key. Then, a data-hider may compress the least significant bits of the encrypted image using a data-hiding key to create a sparse space to accommodate some additional data. With an encrypted image containing additional data, if a receiver has the data-hiding key, he can extract the additional data though he does not know the image content. If the receiver has the encryption key, he can decrypt the received data to obtain an image similar to the original one, but cannot extract the additional data. If the receiver has both the data-hiding key and the encryption key, he can extract the additional data and recover the original content without any error by exploiting the spatial correlation in natural image when the amount of additional data is not too large.

Toward Experiential Mobile Media Processing.

Synopsis:

Smartphone cameras are often used to take pictures and videos because so many people now carry their smartphones everywhere. Therefore, smartphone camera quality plays an important role in overall smartphone evaluation and usage. Although smartphone camera quality has been continually improving, mobile media processing can provide significant

enhancements. The authors discuss mobile media processing techniques, including color and contrast enhancement, backlight image compensation, deblurring, denoising, and video stabilization, within the framework of smartphone camera quality improvement.

Weighted KPCA Degree of Homogeneity Amended NonclassicalReceptive Field Inhibition Model for Salient Contour Extraction in Low-Light-Level Image

Synopsis:

The stimulus response of the classical receptive field (CRF) of neuron in primary visual cortex is affected by its periphery [i.e., non-CRF (nCRF)]. This modulation exerts inhibition, which depends primarily on the correlation of both visual stimulations. The theory of periphery and center interaction with visual characteristics can be applied in night vision information processing. In this paper, a weighted kernel principal component analysis (WKPCA) degree of homogeneity (DH) amended inhibition model inspired by visual perceptual mechanisms is proposed to extract salient contour from complex natural scene in low-light-level image. The core idea is that multifeature analysis can recognize the homogeneity in modulation coverage effectively. Computationally, a novel WKPCA algorithm is presented to eliminate outliers and anomalous distribution in CRF and accomplish principal component analysis precisely. On this basis, a new concept and computational procedure for DH is defined to evaluate the dissimilarity between periphery and center comprehensively. Through amending the inhibition from nCRF to CRF by DH, our model can reduce the interference of noises, suppress details, and textures in homogeneous regions accurately. It helps to further avoid mutual suppression among inhomogeneous regions and contour elements. This paper provides an improved computational visual model with high-performance for contour detection from cluttered natural scene in night vision image.