

INTERNATIONAL JOURNAL OF RESEARCH IN COMMERCE, IT & MANAGEMENT

ijrcm



A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories

Indexed & Listed at:

Ulrich's Periodicals Directory ©, ProQuest, U.S.A., EBSCO Publishing, U.S.A., Cabell's Directories of Publishing Opportunities, U.S.A.

as well as in Open J-Gate, India [link of the same is duly available at infibnet of University Grants Commission (U.G.C.)]

Registered & Listed at: Index Copernicus Publishers Panel, Poland

Circulated all over the world & Google has verified that scholars of more than 1388 Cities in 138 countries/territories are visiting our journal on regular basis.

Ground Floor, Building No. 1041-C-1, Devi Bhawan Bazar, JAGADHRI – 135 003, Yamunanagar, Haryana, INDIA

www.ijrcm.org.in

CONTENTS

Sr. No.	TITLE & NAME OF THE AUTHOR (S)	Page No.
1.	THE IMPACT OF PLANNING AND CONTROL ON SERVICE SMES SUCCESS GAD VITNER & SIBYLLE HEILBRUNN	1
2.	CHALLENGES FOR SMALL AND MEDIUM ENTERPRISES IN INFORMATION TECHNOLOGY IN THE CITY OF BANGALORE, INDIA SULAKSHA NAYAK & DR. HARISHA G. JOSHI	9
3.	ROLE OF MANAGEMENT INFORMATION SYSTEMS IN MANAGERIAL DECISION MAKING OF ORGANIZATIONS IN THE GLOBAL BUSINESS WORLD MD. ZAHIR UDDIN ARIF, MOHAMMAD MIZENUR RAHAMAN & MD. NASIR UDDIN	14
4.	EFFECTS OF CALL CENTER CRM PRACTICES ON EMPLOYEE JOB SATISFACTION DR. ALIYU OLAYEMI ABDULLATEEF	19
5.	DETERMINANTS OF CAPITAL STRUCTURE: EVIDENCE FROM TANZANIA'S LISTED NON FINANCIAL COMPANIES BUNDALA, NTOGWA NG'HABI & DR. CLIFFORD G. MACHOGU	24
6.	RELATIONSHIP BETWEEN INTRINSIC REWARDS AND JOB SATISFACTION: A COMPARATIVE STUDY OF PUBLIC AND PRIVATE ORGANIZATION TAUSIF M.	33
7.	NUCLEAR ENERGY IN INDIA: A COMPULSION FOR THE FUTURE DR. KAMLESH KUMAR DUBEY & SUBODH PANDE	42
8.	CONTEXTUAL FACTORS FOR EFFECTIVE IMPLEMENTATION OF PERFORMANCE APPRAISAL IN THE INDIAN IT SECTOR: AN EMPIRICAL STUDY SUJOYA RAY MOULIK & DR. SITANATH MAZUMDAR	47
9.	A STUDY OF CITIZEN CENTRIC SERVICE DELIVERY THROUGH e-GOVERNANCE: CASE STUDY OF e-MITRA IN JAIPUR DISTRICT RAKESH SINGHAL & DR. JAGDISH PRASAD	53
10.	TWO UNIT COLD STANDBY PRIORITY SYSTEM WITH FAULT DETECTION AND PROVISION OF REST VIKAS SHARMA, J P SINGH JOOREL, RAKESH CHIB & ANKUSH BHARTI	61
11.	MACRO ECONOMIC FACTORS INFLUENCING THE COMMODITY MARKET WITH SPECIAL REFERENCE TO GOLD AND SILVER DR. G. PANDURANGAN, R. MAGENDIRAN, L. S. SRIDHAR & R. RAJKOKILA	68
12.	CRITICAL ANALYSIS OF EXPONENTIAL SMOOTHING METHODS FOR FORECASTING UDAI BHAN TRIVEDI	71
13.	COMPARATIVE STUDY ON RETAIL LIABILITIES, PRODUCTS & SERVICES OF DISTRICT CENTRAL CO-OPERATIVE BANK & AXIS BANK ABHINAV JOG & ZOHRA ZABEEN SABUNWALA	75
14.	SECURE KEY EXCHANGE WITH RANDOM CHALLENGE RESPONSES IN CLOUD BINU V. P & DR. SREEKUMAR A	81
15.	COMPUTATIONAL TRACKING AND MONITORING FOR EFFICIENCY ENHANCEMENT OF SOLAR BASED REFRIGERATION V. SATHYA MOORTHY, P.A. BALAJI, K. VENKAT & G.GOPU	84
16.	FINANCIAL ANALYSIS OF OIL AND PETROLEUM INDUSTRY DR. ASHA SHARMA	90
17.	ANOVA BETWEEN THE STATEMENT REGARDING THE MOBILE BANKING FACILITY AND TYPE OF MOBILE PHONE OWNED: A STUDY WITH REFERENCE TO TENKASI AT VIRUDHUNAGAR DISTRICT DR. S. VALLI DEVASENA	98
18.	VIDEO REGISTRATION BY INTEGRATION OF IMAGE MOTIONS V.FRANCIS DENSIL RAJ & S.SANJEEVE KUMAR	103
19.	ANALYZING THE TRADITIONAL INDUCTION FORMAT AND RE – DESIGNING INDUCTION PROCESS AT TATA CHEMICALS LTD, MITHAPUR PARUL BHATI	112
20.	THE JOURNEY OF E-FILING OF INCOME TAX RETURNS IN INDIA MEENU GUPTA	118
21.	ROLE OF FINANCIAL TECHNOLOGY IN ERADICATION OF FINANCIAL EXCLUSION DR. SARIKA SRIVASTAVA & ANUPAMA AMBUJAKSHAN	122
22.	ATTRITION: THE BIGGEST PROBLEM IN INDIAN IT INDUSTRIES VIDYA SUNIL KADAM	126
23.	INFORMATION TECHNOLOGY IN KNOWLEDGE MANAGEMENT M. SREDEVI	132
24.	A STUDY OF EMPLOYEE ENGAGEMENT & EMPLOYEE CONNECTS' TO GAIN SUSTAINABLE COMPETITIVE ADVANTAGE IN GLOBALIZED ERA NEERU RAGHAV	136
25.	BIG-BOX RETAIL STORE IN INDIA – A CASE STUDY APPROACH WITH WALMART M. P. SUGANYA & DR. R. SHANTHI	142
26.	IMPACT OF INFORMATION TECHNOLOGY ON ORGANISATIONAL CULTURE OF STATE BANK OF INDIA AND ITS ASSOCIATED BANKS IN SRIGANGANAGAR AND HANUMANGARH DISTRICTS OF RAJASTHAN MOHITA	146
27.	USER PERCEPTION TOWARDS WEB, TELEVISION AND RADIO AS ADVERTISING MEDIA: COMPARATIVE STUDY SINDU KOPPA & SHAKEEL AHAMED	149
28.	STUDY OF GROWTH, INSTABILITY AND SUPPLY RESPONSE OF COMMERCIAL CROPS IN PUNJAB: AN ECONOMETRIC ANALYSIS SUMAN PARMAR	156
29.	DEVELOPMENT AND EMPIRICAL VALIDATION OF A LINEAR STYLE PROGRAM ON 'STRUCTURE OF THE CELL' FOR IX GRADE STUDENTS RAMANJEET KAUR	160
30.	PERFORMANCE APPRAISAL OF INDIAN BANKING SECTOR: A COMPARATIVE STUDY OF SELECTED PUBLIC AND FOREIGN BANKS SAHILA CHAUDHRY	163
	REQUEST FOR FEEDBACK	173

CHIEF PATRON

PROF. K. K. AGGARWAL

Chancellor, Lingaya's University, Delhi
Founder Vice-Chancellor, Guru Gobind Singh Indraprastha University, Delhi
Ex. Pro Vice-Chancellor, Guru Jambheshwar University, Hisar

PATRON

SH. RAM BHAJAN AGGARWAL

Ex. State Minister for Home & Tourism, Government of Haryana
Vice-President, Dadri Education Society, Charkhi Dadri
President, Chinar Syntex Ltd. (Textile Mills), Bhiwani

CO-ORDINATOR

AMITA

Faculty, Government M. S., Mohali

ADVISORS

DR. PRIYA RANJAN TRIVEDI

Chancellor, The Global Open University, Nagaland

PROF. M. S. SENAM RAJU

Director A. C. D., School of Management Studies, I.G.N.O.U., New Delhi

PROF. M. N. SHARMA

Chairman, M.B.A., Haryana College of Technology & Management, Kaithal

PROF. S. L. MAHANDRU

Principal (Retd.), Maharaja Agrasen College, Jagadhri

EDITOR

PROF. R. K. SHARMA

Professor, Bharti Vidyapeeth University Institute of Management & Research, New Delhi

CO-EDITOR

DR. BHAVET

Faculty, M. M. Institute of Management, Maharishi Markandeshwar University, Mullana, Ambala, Haryana

EDITORIAL ADVISORY BOARD

DR. RAJESH MODI

Faculty, Yanbu Industrial College, Kingdom of Saudi Arabia

PROF. SANJIV MITTAL

University School of Management Studies, Guru Gobind Singh I. P. University, Delhi

PROF. ANIL K. SAINI

Chairperson (CRC), Guru Gobind Singh I. P. University, Delhi

DR. SAMBHAVNA

Faculty, I.I.T.M., Delhi

DR. MOHENDER KUMAR GUPTA

Associate Professor, P. J. L. N. Government College, Faridabad

DR. SHIVAKUMAR DEENE

Asst. Professor, Dept. of Commerce, School of Business Studies, Central University of Karnataka, Gulbarga

MOHITA

Faculty, Yamuna Institute of Engineering & Technology, Village Gadholi, P. O. Gadholi, Yamunanagar

ASSOCIATE EDITORS

PROF. NAWAB ALI KHAN

Department of Commerce, Aligarh Muslim University, Aligarh, U.P.

PROF. ABHAY BANSAL

Head, Department of Information Technology, Amity School of Engineering & Technology, Amity University, Noida

PROF. A. SURYANARAYANA

Department of Business Management, Osmania University, Hyderabad

DR. ASHOK KUMAR

Head, Department of Electronics, D. A. V. College (Lahore), Ambala City

DR. SAMBHAV GARG

Faculty, M. M. Institute of Management, Maharishi Markandeshwar University, Mullana, Ambala, Haryana

PROF. V. SELVAM

SSL, VIT University, Vellore

DR. PARDEEP AHLAWAT

Reader, Institute of Management Studies & Research, Maharshi Dayanand University, Rohtak

S. TABASSUM SULTANA

Associate Professor, Department of Business Management, Matrusri Institute of P.G. Studies, Hyderabad

SURJEET SINGH

Asst. Professor, Department of Computer Science, G. M. N. (P.G.) College, Ambala Cantt.

TECHNICAL ADVISOR

AMITA

Faculty, Government H. S., Mohali

MOHITA

Faculty, Yamuna Institute of Engineering & Technology, Village Gadholi, P. O. Gadholi, Yamunanagar

FINANCIAL ADVISORS

DICKIN GOYAL

Advocate & Tax Adviser, Panchkula

NEENA

Investment Consultant, Chambaghat, Solan, Himachal Pradesh

LEGAL ADVISORS

JITENDER S. CHAHAL

Advocate, Punjab & Haryana High Court, Chandigarh U.T.

CHANDER BHUSHAN SHARMA

Advocate & Consultant, District Courts, Yamunanagar at Jagadhri

SUPERINTENDENT

SURENDER KUMAR POONIA

CALL FOR MANUSCRIPTS

We invite unpublished novel, original, empirical and high quality research work pertaining to recent developments & practices in the area of Computer, Business, Finance, Marketing, Human Resource Management, General Management, Banking, Insurance, Corporate Governance and emerging paradigms in allied subjects like Accounting Education; Accounting Information Systems; Accounting Theory & Practice; Auditing; Behavioral Accounting; Behavioral Economics; Corporate Finance; Cost Accounting; Econometrics; Economic Development; Economic History; Financial Institutions & Markets; Financial Services; Fiscal Policy; Government & Non Profit Accounting; Industrial Organization; International Economics & Trade; International Finance; Macro Economics; Micro Economics; Monetary Policy; Portfolio & Security Analysis; Public Policy Economics; Real Estate; Regional Economics; Tax Accounting; Advertising & Promotion Management; Business Education; Management Information Systems (MIS); Business Law, Public Responsibility & Ethics; Communication; Direct Marketing; E-Commerce; Global Business; Health Care Administration; Labor Relations & Human Resource Management; Marketing Research; Marketing Theory & Applications; Non-Profit Organizations; Office Administration/Management; Operations Research/Statistics; Organizational Behavior & Theory; Organizational Development; Production/Operations; Public Administration; Purchasing/Materials Management; Retailing; Sales/Selling; Services; Small Business Entrepreneurship; Strategic Management Policy; Technology/Innovation; Tourism, Hospitality & Leisure; Transportation/Physical Distribution; Algorithms; Artificial Intelligence; Compilers & Translation; Computer Aided Design (CAD); Computer Aided Manufacturing; Computer Graphics; Computer Organization & Architecture; Database Structures & Systems; Digital Logic; Discrete Structures; Internet; Management Information Systems; Modeling & Simulation; Multimedia; Neural Systems/Neural Networks; Numerical Analysis/Scientific Computing; Object Oriented Programming; Operating Systems; Programming Languages; Robotics; Symbolic & Formal Logic and Web Design. The above mentioned tracks are only indicative, and not exhaustive.

Anybody can submit the soft copy of his/her manuscript **anytime** in M.S. Word format after preparing the same as per our submission guidelines duly available on our website under the heading guidelines for submission, at the email addresses: infoijrcm@gmail.com or info@ijrcm.org.in.

GUIDELINES FOR SUBMISSION OF MANUSCRIPT

1. COVERING LETTER FOR SUBMISSION:

DATED: _____

THE EDITOR
IJRCM

Subject: SUBMISSION OF MANUSCRIPT IN THE AREA OF _____.

(e.g. Finance/Marketing/HRM/General Management/Economics/Psychology/Law/Computer/IT/Engineering/Mathematics/other. **please specify**)

DEAR SIR/MADAM

Please find my submission of manuscript entitled ' _____ ' for possible publication in your journals.

I hereby affirm that the contents of this manuscript are original. Furthermore, it has neither been published elsewhere in any language fully or partly, nor is it under review for publication elsewhere.

I affirm that all the author (s) have seen and agreed to the submitted version of the manuscript and their inclusion of name (s) as co-author (s).

Also, if my/our manuscript is accepted, I/We agree to comply with the formalities as given on the website of the journal & you are free to publish our contribution in any of your journals.

NAME OF CORRESPONDING AUTHOR:

Designation:

Affiliation with full address, contact numbers & Pin Code:

Residential address with Pin Code:

Mobile Number (s):

Landline Number (s):

E-mail Address:

Alternate E-mail Address:

NOTES:

- a) The whole manuscript is required to be in **ONE MS WORD FILE** only (pdf. version is liable to be rejected without any consideration), which will start from the covering letter, inside the manuscript.
- b) The sender is required to mention the following in the **SUBJECT COLUMN** of the mail:
New Manuscript for Review in the area of (Finance/Marketing/HRM/General Management/Economics/Psychology/Law/Computer/IT/Engineering/Mathematics/other, please specify)
- c) There is no need to give any text in the body of mail, except the cases where the author wishes to give any specific message w.r.t. to the manuscript.
- d) The total size of the file containing the manuscript is required to be below **500 KB**.
- e) Abstract alone will not be considered for review, and the author is required to submit the complete manuscript in the first instance.
- f) The journal gives acknowledgement w.r.t. the receipt of every email and in case of non-receipt of acknowledgment from the journal, w.r.t. the submission of manuscript, within two days of submission, the corresponding author is required to demand for the same by sending separate mail to the journal.

2. **MANUSCRIPT TITLE:** The title of the paper should be in a 12 point Calibri Font. It should be bold typed, centered and fully capitalised.

3. **AUTHOR NAME (S) & AFFILIATIONS:** The author (s) **full name, designation, affiliation (s), address, mobile/landline numbers, and email/alternate email address** should be in italic & 11-point Calibri Font. It must be centered underneath the title.

4. **ABSTRACT:** Abstract should be in fully italicized text, not exceeding 250 words. The abstract must be informative and explain the background, aims, methods, results & conclusion in a single para. Abbreviations must be mentioned in full.

5. **KEYWORDS:** Abstract must be followed by a list of keywords, subject to the maximum of five. These should be arranged in alphabetic order separated by commas and full stops at the end.
6. **MANUSCRIPT:** Manuscript must be in **BRITISH ENGLISH** prepared on a standard A4 size **PORTRAIT SETTING PAPER**. It must be prepared on a single space and single column with 1" margin set for top, bottom, left and right. It should be typed in 8 point Calibri Font with page numbers at the bottom and centre of every page. It should be free from grammatical, spelling and punctuation errors and must be thoroughly edited.
7. **HEADINGS:** All the headings should be in a 10 point Calibri Font. These must be bold-faced, aligned left and fully capitalised. Leave a blank line before each heading.
8. **SUB-HEADINGS:** All the sub-headings should be in a 8 point Calibri Font. These must be bold-faced, aligned left and fully capitalised.
9. **MAIN TEXT:** The main text should follow the following sequence:

INTRODUCTION

REVIEW OF LITERATURE

NEED/IMPORTANCE OF THE STUDY

STATEMENT OF THE PROBLEM

OBJECTIVES

HYPOTHESES

RESEARCH METHODOLOGY

RESULTS & DISCUSSION

FINDINGS

RECOMMENDATIONS/SUGGESTIONS

CONCLUSIONS

SCOPE FOR FURTHER RESEARCH

ACKNOWLEDGMENTS

REFERENCES

APPENDIX/ANNEXURE

It should be in a 8 point Calibri Font, single spaced and justified. The manuscript should preferably not exceed **5000 WORDS**.

10. **FIGURES & TABLES:** These should be simple, crystal clear, centered, separately numbered & self explained, and **titles must be above the table/figure**. **Sources of data should be mentioned below the table/figure**. It should be ensured that the tables/figures are referred to from the main text.
11. **EQUATIONS:** These should be consecutively numbered in parentheses, horizontally centered with equation number placed at the right.
12. **REFERENCES:** The list of all references should be alphabetically arranged. The author (s) should mention only the actually utilised references in the preparation of manuscript and they are supposed to follow **Harvard Style of Referencing**. The author (s) are supposed to follow the references as per the following:
 - All works cited in the text (including sources for tables and figures) should be listed alphabetically.
 - Use (ed.) for one editor, and (ed.s) for multiple editors.
 - When listing two or more works by one author, use --- (20xx), such as after Kohl (1997), use --- (2001), etc, in chronologically ascending order.
 - Indicate (opening and closing) page numbers for articles in journals and for chapters in books.
 - The title of books and journals should be in italics. Double quotation marks are used for titles of journal articles, book chapters, dissertations, reports, working papers, unpublished material, etc.
 - For titles in a language other than English, provide an English translation in parentheses.
 - The location of endnotes within the text should be indicated by superscript numbers.

PLEASE USE THE FOLLOWING FOR STYLE AND PUNCTUATION IN REFERENCES:

BOOKS

- Bowersox, Donald J., Closs, David J., (1996), "Logistical Management." Tata McGraw, Hill, New Delhi.
- Hunker, H.L. and A.J. Wright (1963), "Factors of Industrial Location in Ohio" Ohio State University, Nigeria.

CONTRIBUTIONS TO BOOKS

- Sharma T., Kwatra, G. (2008) Effectiveness of Social Advertising: A Study of Selected Campaigns, Corporate Social Responsibility, Edited by David Crowther & Nicholas Capaldi, Ashgate Research Companion to Corporate Social Responsibility, Chapter 15, pp 287-303.

JOURNAL AND OTHER ARTICLES

- Schemenner, R.W., Huber, J.C. and Cook, R.L. (1987), "Geographic Differences and the Location of New Manufacturing Facilities," Journal of Urban Economics, Vol. 21, No. 1, pp. 83-104.

CONFERENCE PAPERS

- Garg, Sambhav (2011): "Business Ethics" Paper presented at the Annual International Conference for the All India Management Association, New Delhi, India, 19–22 June.

UNPUBLISHED DISSERTATIONS AND THESES

- Kumar S. (2011): "Customer Value: A Comparative Study of Rural and Urban Customers," Thesis, Kurukshetra University, Kurukshetra.

ONLINE RESOURCES

- Always indicate the date that the source was accessed, as online resources are frequently updated or removed.

WEBSITE

- Garg, Bhavet (2011): Towards a New Natural Gas Policy, Political Weekly, Viewed on January 01, 2012 <http://epw.in/user/viewabstract.jsp>

VIDEO REGISTRATION BY INTEGRATION OF IMAGE MOTIONS

V.FRANCIS DENSIL RAJ
ASST. PROFESSOR
DEPARTMENT OF MCA
ANNA UNIVERSITY OF TECHNOLOGY
MADURAI

S.SANJEEVE KUMAR
ASST. PROFESSOR
DEPARTMENT OF CSE
ANNA UNIVERSITY OF TECHNOLOGY
MADURAI

ABSTRACT

In this research, we consider the problems of registering multiple video sequences dynamic scenes which are not limited non rigid objects such as fireworks, blasting, high speed car moving taken from different vantage points. In this paper we propose a simple algorithm we can create different frames on particular videos moving for matching such complex scenes. Our algorithm does not require the cameras to be synchronized, and is not based on frame-by-frame or volume-by-volume registration. Instead, we model each video as the output of a linear dynamical system and transform the task of registering the video sequences to that of registering the parameters of the corresponding dynamical models. In this paper we use of a joint frame together to form distinct frame concurrently. The joint identification and the Jordan canonical form are not only applicable to the case of registering video sequences, but also to the entire genre of algorithms based on the dynamic texture model. We have also shown that out of all the possible choices for the method of identification and canonical form, the JID using JCF performs the best.

KEYWORDS

Dynamic textures, non-rigid dynamical scenes, video registration.

INTRODUCTION

The purpose of this paper is to outline a technique for finding the registration between two frames from a sequence of video that corresponds to the camera motion that also provides a means of detecting, and approximately segmenting, the moving objects within that scene. The approach taken is to split the image up into a series of blocks, and run a search comparing different registration parameters to find the transform that most effectively maps the motion of the scene background by means of grouping. A sequence of video generally involves two main unknowns: the movement of the camera, and the form and motion of the moving objects in the scene. This can involve many different motions, all of which must be detected and calculated separately, which is a very difficult task.

Image registration is the process of matching a pair of similar images in terms of the rotation, translation, scale and shear required to make those images correctly align. The particular focus of this paper is the application of this on two frames from a video sequence to describe image motion. Applications include automated image and video in painting for special effects in the film industry, use with mobile security systems, and vision systems for autonomous agents. For images that do not contain moving objects, the process of extracting information about the movement of the camera for the scene is relatively straightforward – by using an affine transformation model it is simply a matter of finding transformation parameters (rotation and translation) that match the two images using a simple error metric like sum of squares difference. However, once moving objects are introduced to the scene, the registration process becomes more complicated. For a situation where the motion of the camera is unknown, and the location of any moving objects is unknown, a registration algorithm has a significant amount of information that it needs to infer. Different sections of the images will have different motions (for example the background moving in one direction due to camera motion, while a person walking through the scene moves in a different direction). These areas of separate motion need to be identified and the magnitude and direction of the motion needs to be independently calculated.

This paper looks at an approach to this problem using a block-based technique for the registration. This enables differentiation between the dominant object in the scene (assumed to be the background) and other objects, which move relative to the dominant object. This provides information on the motion of the camera filming the scene. The task of identifying moving objects within a scene is another area of computer vision that is attracting attention. This paper proposes to use the information obtained through the process of registering the two images to provide a reasonable identification of the object location, for eventual application in the area of video in painting.

METHOD**AFFINE REGISTRATION**

The technique used for registration in this paper is parametric affine registration. We currently restrict the algorithm from full affine to a three-parameter model, using rotation and x- and y- translation. The decision to ignore shear and scale considers that the effect of these between adjacent frames in standard video is likely to be minimal, and the block-based approach should allow for this to be detected without explicitly searching for it. Though it may have some impact on accuracy, the gain in efficiency obtained by ignoring scale and shear is significant.

The three-parameter affine transformation model for mapping the registration of a single pixel between two images

(I_0 and I_1) is as follows

$$fx_0, y_0 = x_1 y_1 = \cos(\theta) \quad \sin(\theta) - \sin(\theta) \cos(\theta) + x_0 y_0 + t_x t_y$$

where $x_0 y_0$ is the position of a pixel in the original image

I_0 and $x_1 y_1$ is the position of the corresponding pixel in the second image, I_1 .

The rotation is conducted about the point (0, 0) on the image (the top left corner), so for a different centre of rotation the image must be offset. The values t_x and t_y are the x- and y-translations respectively. The registration itself is a minimisation problem: identifying values for t_x , t_y , and θ that minimise an error value between image I_1 and the transformed I_0 . The error used for this paper is the sum-of-squares difference between the RGB colour values of each pixel over the image or image section:

$$E = \sum (I_0(x_i, y_i) - I_1(f(x_i, y_i)))^2$$

As algorithmic efficiency is a major concern, the use of a gradient descent algorithm to optimise the search was investigated. For this application, however, gradient descent is not ideal – the frequent occurrence of local minima throughout the search space, a high susceptibility to noise, and an inability to effectively process areas of low colour differential severely hamper its effectiveness. The efficiency problems can be greatly reduced with an approach using a series of

iterative searches over reduced scale images. The scale of the reduction required depends on the image sizes involved – for the relatively small samples analysed in this project, an initial coarse search is conducted over a quarter-scale image, which is then refined over a half scale image before a full-scale image is used to obtain the sub-pixel accuracy required.

THE INTEGRATION APPROACH

Attempting to match a transform to the entirety of an image is impractical for a registration technique on video containing moving objects, as different parts of the image will be moving different amounts. Therefore, it makes sense to acknowledge that there will be objects, and allow them to be excluded from the registration. Methods like optical flow can achieve this by mapping pixels individually, but individual pixels are highly susceptible to noise, which can in turn affect the registration.

If working over the entire image is too broad, and looking at individual pixels is too fine, then it is logical to try and find a sort of middle ground. By splitting the image up into a series of blocks and tracking these separately, errors caused by certain types of noise (most notably, the blurring that is a common result of video compression) can be minimized, and sections of the image containing moving objects can be eliminated from the registration problem.

There are two approaches to the block-based image registration. The first involves running the registration search algorithm for each block individually to calculate the best transform for each block. These results can then be correlated in the form of a three dimensional histogram, where the largest group corresponds to the transformation that describes the motion of the largest 'object', which can be taken to be the background. This method is inherently slow, as the search space is fairly large. For this reason, an iterative search approach over a multi-resolution pyramid is employed. The algorithm is initially run at quarter-scale or smaller (depending on the size of the original image) with broad search parameters, then the results are aggregated and used to determine a much smaller search area, which is then applied to a half-scale image. The search parameters are then refined again, before finally being applied to the original full-scale image. Susceptibility to noise and difficulty distinguishing large areas of colour lead to a proportion of blocks returning incorrect results for the registration. For this reason, at each stage of the refinement the overall best result is used as the base for the next iteration of the registration for each block. The second method is considerably faster, but also introduces a much greater potential for error. It involves picking a block at random and running the affine search over that sole block to find the transformation that fits that block to the next image. That transformation is then applied to all the blocks and thresholding is used to identify which blocks agree with the transform, which in turn gives the approximate area of the object whose movement is mapped by the transform.

Largely, this causes problems because the system is based on arbitrary thresholds. If the threshold identifying which blocks agree with a transform is too low, blocks which do agree may be excluded, resulting in a lower value for object area, which could result in the background being disregarded as being just a small object. If the threshold is too high, the algorithm may find local minima that is deemed an acceptable transform, but is not the global minima – most likely to occur when the sampled block contains parts of more than one object. The other threshold that may cause problems is the one that is used to determine whether an object is large enough to be identified as the background. If the value is too high then there may be no objects that meet the criteria and the algorithm will never complete, or, conversely, if it is set too low then a smaller object may be incorrectly identified as the background. This particular problem may be countered by eliminating blocks that have already been grouped and resampling until every region of the image has been identified, at which point the largest of these can be positively identified as the background. This will, however, significantly increase the computational cost of the algorithm (as it will no longer halt as soon as it finds an acceptably large object) as well as introducing the possibility of error when a block is assigned into a group that 'fits' but is not the optimum grouping for the block. In theory, this algorithm appears to solve the problem that is the target of this paper: identifying and tracking each individual object in the scene. In practice, however, this is unachievable. Firstly, there are the blocks that contain parts of more than one object; these will seldom match the transform of any object, so will not be grouped correctly. Similarly, blocks that are resampled to by the algorithm and that exist in areas of solid colour (i.e. have no distinguishing features/colours) will likely not track correctly, which would also interfere with the results. Figure 1 shows the chosen transformations for a registration search run over an image pair using 20x20 blocks. Only a single search is run on the full-scale image, to demonstrate the registration process without the corrective measures that are used in the iterative coarse-to-fine search. The number of blocks that produce conflicting registrations demonstrates the high occurrence of local minima over the search, and thus the necessity of the corrective groupings used in the iterative search.

FIGURE 1



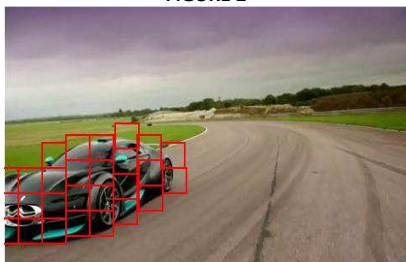
This shows the detected motion paths of blocks across the image after one iteration of the affine search on the full-scale image. (See Figure 3 for the images used in this registration). The necessity of an iterative search approach is shown by the number of apparently random motion paths that are produced.

EXTRACTING OBJECT INFORMATION

The extraction of object information is based on the premise that most of the image is background, which moves according to the motion of the camera, with certain sections of the image not conforming to this motion – those sections belonging to independently moving objects. During the process of tracking the camera motion, the image sections that do not conform to the global motion must be detected and eliminated in order for the result to be accurate. Thus, the location and, to a certain extent, the motion of these objects is already available once the camera tracking algorithm has run. Due to the necessarily coarse nature of the block-based approach, an exact outline of the objects is not provided (see Figure 2), however enough information is discovered to make the task of accurately segmenting these objects much easier with the application of an algorithm such as active contours.

For the purpose of comparison, a second method based on a difference algorithm applied post-registration is used. The algorithm uses the existing block structure to reduce the impact of noise on the results. Each block is marked as belonging to an independently moving object if sufficient pixels within the block are significantly different. The level of difference that is taken as significant is determined by an adjustable threshold, which requires changes depending on the level of contrast. Figure 2 shows an example of detection of a moving object using the block-based method. The image was first registered to a second image, then the block information was used to highlight moving objects within the scene. The blocks shown on the image indicate the detected object area between background and foreground in the samples.

FIGURE 2



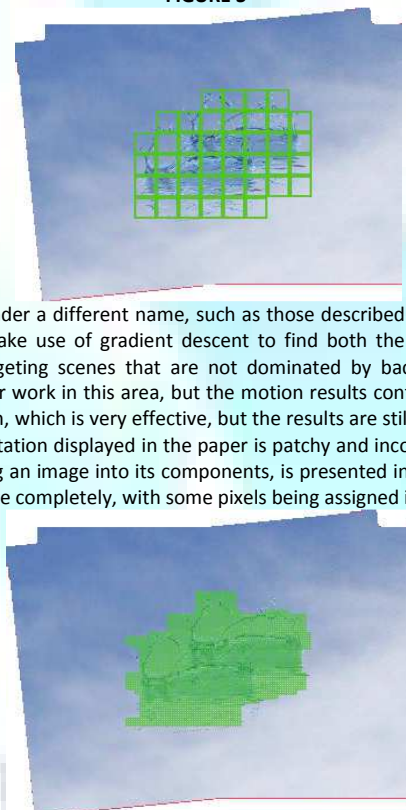
Example of detection of a moving object using the block-based method.

The image was first registered to a second image, and then the block information was used to highlight moving objects within the scene. The blocks shown on the image indicate the detected object area. The method using registration information to identify moving objects works well in a broader range of situations. While the difference algorithm is often slightly more effective in high contrast scenarios, it struggles with low contrast.

RELATED WORK

There are many methods of affine registration; good overviews of the work in the area are offered by [1], [2]. A common application makes use of registration techniques for mapping medical imaging scans to assist in the detection of anomalies, however the types of deformation involved do not transfer to the area of camera motion effectively. Patch-tracking is one area within the scope of registration and object tracking that has been pursued, with two different interpretations of its meaning available. The patch-tracking algorithm put forward in [3] is focused on tracking an object, and requires prior knowledge about the form of that object, which is not particularly useful here. [4] Offers an entirely different patch-based algorithm, which uses patches similar to the blocks used in this paper to solve regression equations for image distortion. The results presented here are impressive; however it does not deal with the existence of moving objects in the scene.

FIGURE 3



Window-based algorithms offer a similar sort of idea under a different name, such as those described in [5], [6]. These algorithms are designed to identify the camera motion in scenes with moving objects, and make use of gradient descent to find both the global minimum and a secondary, local minimum that describes the camera motion – this is specifically targeting scenes that are not dominated by background. In terms of a registration approach to object segmentation, [7] establishes an effective foundation for work in this area, but the motion results contain some sections of image unrelated to the objects. [8] Presents with a more complete example of this approach, which is very effective, but the results are still influenced by outliers in the sequence. [9] Also provides some work in this area, but the resulting object segmentation displayed in the paper is patchy and incomplete. Another approach to the problem, using a pixel-wise optical flow method for registering and segmenting an image into its components, is presented in [10]. It produces some impressive results, but due to its pixel-wise nature the layers of the image will not separate completely, with some pixels being assigned incorrectly.

The work presented in [7] is somewhat in between. A block based approach is used for the object detection stage but not for the global motion estimation, and the object detection uses a difference based approach rather than using the information available from the registration.

MULTI INTEGRATION MAPPING

By introducing extra blocks across the image that overlap but are offset from the main registration blocks before the final fine-search of the registration, a higher degree of accuracy can be achieved for the object tracking at a certain cost in terms of computing power. These extra blocks are registered in the same way as the original blocks, and the results are aggregated using the logical AND operator to produce a finer resolution for the tracking result – only areas that are marked as moving objects by *all* of the blocks that contain them are included in the tracking result. This can be seen as a basic form of image super-resolution. For example, using one extra set of pixel blocks, offset from the original grid by 10 pixels in both the x- and y-directions, would produce a tracking result constructed of 5×5 pixel blocks rather than the 20×20 blocks that would otherwise be the result.

Figure 3 shows both the basic approach and the multi-block approach (using three extra block sets over the image). The multi-block approach provides a segmentation that does not include as much non-object data, but it still loses small sections of the edge of the object.

Figure 3 The first image shows the object detection using the basic approach, with the detected object area shown by the registration blocks it is detected in.

The second image shows the multi-block approach, this time the detected area is shown as a grid to simplify the drawing algorithm. The images that the object detection is shown on is the mean image taken after registration (allowing the position of the object in both frames to be seen on the same picture).

DYNAMIC TEXTURES FRAMEWORK

Given a video sequence $\{I(t)\}_{t=1}^F$, we model the temporal evolution of its intensities as the output of an LDS. The equations that model the sequence are given by

$$Z(t+1) = AZ(t) + BV(t), \quad (1)$$

$$I(t) = C^0 + CZ(t) + W(t). \quad (2)$$

The parameters of this model can be classified into three types, namely, the appearance, dynamics, and noise parameters. The vector $z(t) \in \mathbb{R}^n$ represents the hidden state of the system at time t . Its evolution is controlled by the dynamics matrix $A \in \mathbb{R}^{n \times n}$ and the input-to-state matrix $B \in \mathbb{R}^{n \times q}$. These parameters are termed the dynamics parameters of the dynamic texture model. The parameter $C \in \mathbb{R}^{p \times n}$ maps the hidden state to the image and the vector $C^0 \in \mathbb{R}^p$ is the temporal mean of the video sequence.

These parameters are called the appearance parameters of the dynamic texture model. The noise parameters are given by the zero-mean Gaussian processes $v(t) \sim N(0, Q)$ and $w(t) \sim N(0, R)$, which model the process noise and the measurement noise, respectively. The order of the system is given by n and p is the number of pixels in the image. The advantage of using this model is that it enables us to decouple the appearance parameters of the video sequence from the dynamics parameters. Thus, if one is interested in recovering appearance-based information from the video sequence, such as optical flow or, in our case, the spatial registration, then one only needs to deal with the appearance parameters. This allows us to recover the spatial registration independent of the temporal alignment of the sequence, as will be seen in the next section.

RECOVERING THE SPATIAL-TEMPORAL TRANSFORMATION FROM THE PARAMETERS OF DYNAMIC TEXTURES

As motivated in the previous section, spatial registration can be recovered using a subset of the parameters of the LDS. In this section, we explore the relationship between the parameters of two video sequences that are spatially and temporally transformed versions of each other.

Let $x = (x, y)$ be the coordinates of a pixel in the image. We define $C^i(x)$ to be the i th column of the C matrix reshaped as an image. Likewise, we define $C^0(x)$ to be the mean of the video sequence reshaped as an image. With this notation, the dynamic texture model can be rewritten as

$$I(x, t) = \sum_{i=0}^n z^i(t) C^i(x) + w(t), \quad (3)$$

Where $z^0(t) = 1$. Therefore, under the dynamic texture model, a video sequence is interpreted as an affine combination of n basis images and the mean image. We call these $n + 1$ images $\{C^i(x)\}_{i=0}^n$ the dynamic appearance images. In the following analysis, we consider a video sequence and its corresponding LDS. We show that a spatial and temporal transformation on the video sequences induces a spatial and temporal transformation in the parameters of the LDS. We first consider the simple case of the two video sequences being just spatially transformed version of each other. Such video sequences are termed as synchronized video sequences. We then consider the more general case of two video sequences being both spatially and temporally transformed versions of each other. Such video sequences are also known as unsynchronized video sequences.

SYNCHRONIZED VIDEO SEQUENCES

Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be any spatial transformation relating the frames from each video, such as a 2D affine transformation or a homography. The relationship between two synchronized video sequence is then given by $\tilde{I}(x, t) = I(T(x), t)$. Consider now the following LDSs, with the evolution of the hidden state as

$$z(t+1) = Az(t) + Bv(t), \quad (4)$$

and the outputs defined as

$$I(x, t) = \sum_{i=0}^n z^i(t) C^i(x) + w(t), \quad (5)$$

$$\tilde{I}(x, t) = \sum_{i=0}^n z^i(t) C^i(T(x)) + W(t). \quad (6)$$

We can see that $\tilde{I}(x, t) = I(T(x), t)$. This shows that when a constant spatial transformation is applied to all the frames in a video sequence, the transformed video can be represented by an LDS that has the same A and B matrices as the original video. The main difference is that the dynamic appearance images $\{C^i(x)\}_{i=0}^n$ are transformed by the same spatial transformation applied to the frames of the sequences, i.e., $\{C^i(x) = C^i(T(x))\}_{i=0}^n$.

UNSYNCHRONIZED VIDEO SEQUENCES

In this case, in addition to the spatial transformation, we now introduce a temporal lag between the two video sequences denoted by T . The relationship between two unsynchronized video sequence can be represented as $\tilde{I}(x, t) = I(T(x), t + T)$. Now let us consider the following two systems, with the evolution of the hidden states given by (4), and the outputs defined as

$$I(x, t) = \sum_{i=0}^n z^i(t) C^i(x) + w(t), \quad (7)$$

$$\tilde{I}(x, t) = \sum_{i=0}^n z^i(t+T) C^i(T(x)) + W(t+T). \quad (8)$$

We now see that the above equations model two unsynchronized sequences. Thus, a video sequence that is a spatially and temporally transformed version of the original video sequence can be represented with an LDS with the same A and the same B as the original video sequences. However, in addition to the C matrix being modified by the spatial transformation, as in the synchronized case, we also have a different initial state. Instead of the video sequence starting at $z(0)$, the initial state now is $z(T)$. Nevertheless, if one wants to only recover the spatial transformation, the C matrices of the two LDSs are the only parameters that need to be compared. Thus, given two video sequences, either synchronized or the unsynchronized, in order to recover the spatial registration, we only need to compare the C matrices. But this is under the assumption that both the A matrix and the state of the system $z(t)$, modulo a temporal shift $T_i \in \mathbb{Z}$, for the two systems remain the same. The rationale behind this assumption is that since the video sequences are of the same scene, the evolution of the hidden states remains the same. More specifically, the objects in the scene undergo the same deformation; hence, they possess the same dynamics. However, if one learns the parameters of the LDSs from the data using existing methods, one encounter two problems. The first problem is that the A matrix is not the same for the different LDSs. Second, the C matrices that are recovered are only unique up to an invertible transformation. Hence, in order to perform the registration, we address these issues in the next section.

RECOVERING TRANSFORMATION PARAMETERS FROM THE DYNAMIC TEXTURE MODEL

In the previous section, we have introduced the dynamic texture model and shown how the model parameters vary for video sequences taken at different viewpoints and time instances. In this section, we will show how the spatial transformation can be recovered using the parameters of the LDSs. In Section 3.4.1, we review the classical system identification algorithm for learning the parameters of an LDS and show that the recovered parameters are not unique. Since we would like to compare the parameters of different LDSs, our first task is to remove such ambiguities. This issue is addressed in Section 3.4.2. We then, in Section 3.4.3, show how we can enforce the dynamics of multiple video sequences to be the same. Finally, in Section 3.4.4, we propose an algorithm to recover the spatial and temporal transformation from the dynamic appearance images of two video sequences.

PARAMETER IDENTIFICATION

Given a video sequence $\{I(t)\}_{t=1}^F$, the first step is to identify the parameters of the LDS. There are several choices for the identification of such systems from the classical system identification literature, e.g., subspace identification methods such as N4SID [14]. The problem with such methods is that as the size of the output increases, these methods become computationally very expensive. Hence, traditionally, the method of identification for dynamic textures has been a suboptimal solution proposed in [12]. This method is essentially a Principal Component Analysis (PCA) decomposition of the video sequence.

$\{I(t)\}_{t=1}^F$ the mean $C^0 = 1/F \sum_{t=1}^F I(t)$ is first calculated. The parameters of the system are then identified from the compact (rank- n) SVD of the mean subtracted data matrix as

$$[I(1) - C^0, \dots, I(F) - C^0] = U(SV)^T = CZ, \quad (9)$$

where $Z = [Z(1) \dots Z(F)]$. Given Z , the parameter A is obtained as the least-square solution to the system of linear equation $A[Z(1) \dots Z(F-1)] = [Z(2) \dots Z(F)]$.

It is well known that the factorization obtained from the SVD is unique up to an invertible transformation, i.e., the factors that are recovered are (CP^{-1}, P^T) , where $P \in \mathbb{R}^{n \times n}$ is an arbitrary invertible matrix. Hence, the LDSs with (A, B, C) and (PAP^{-1}, PB, CP^{-1}) both generate the same output process. This fact does not pose a problem when dealing with a single video sequence. However, when one wants to compare the parameters identified from multiple sequences, each set of identified parameters could potentially be computed with respect to a different basis. Since our goal is to compare the C matrices, to perform the registration we need to ensure that different C matrices are in the same basis. In order to address this issue, in the next section, we outline a method to account for the basis change. We propose to do this by using a canonical form and converting the parameters into the canonical form.

CANONICAL FORMS FOR PARAMETER COMPARISON

Given the parameters of an LDS (A, B, C) , the family of parameters that generate the same output process is given by (PAP^{-1}, PB, CP^{-1}) . There are several approaches to removing the ambiguities from the system parameters. For instance, one can restrict the columns of the C matrix to be orthogonal. Exploiting this fact, one option to overcome the basis ambiguities is to project all the C matrices into the subspace spanned by one of the C matrices. In [11], Chan and Vasconcelos used such an approach where one sequence was chosen as the reference and the parameters of the other sequence were converted into the basis of the reference sequence. One drawback of such a method is that it requires a reference sequence. Choosing such a reference sequence might not always be feasible. An alternate approach from linear systems theory, to address the basis issue, is to use a canonical form. The advantage of such canonical forms is that the model parameters in the canonical forms have a specific structure. As a consequence, if the model parameters identified using the suboptimal approach is converted to the canonical form, the parameters are in the same basis. This removes the basis ambiguity induced in the suboptimal identification algorithm due to the SVD factorization. Also, using the canonical form does not require a reference sequence.

If one refers to the literature from linear systems theory, several canonical forms have been proposed for the model parameters of an LDS in the particular case of a single output system, i.e., $p = 1$. Although, in principle, any canonical form can be used to overcome the ambiguities, the fact that these LDSs model the temporal evolution of the intensities of pixels of a video sequence poses some constraints in the choice of the canonical form. For example, we do not want such forms to be complex. This would make it difficult to perform comparison between parameters of different systems. Also, even though, theoretically, all of the canonical forms are equivalent, in practice they differ in their numerical stability. Vidal and Ravichandran in [17] used a diagonal form for the A matrix. Since the equivalence class of parameters for A is PAP^{-1} , i.e., a similarity transformation, the diagonal form reduces to the diagonal matrix of eigenvalues of A. Thus, the resulting parameters in canonical form can be complex, since the eigenvalues of A can be complex. To overcome this, the Reachability Canonical Form (RCF) was used in [15]. The RCF is given by

$$A_c = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 0 \end{bmatrix}, \quad B_c = \begin{bmatrix} -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}^T \in \mathbb{R}^{n \times 1}, \quad (10)$$

where $A^n + a_{n-1}A^{n-1} + \dots + a_0I = 0$ is the characteristic polynomial of A and I_{n-1} is the identity matrix of size $n-1$. The problem with the RCF is that it uses the pair (A, B) to convert the system into canonical form. For most common applications of dynamic textures, such as registration and recognition, it is preferable to have a canonical form based on the parameters (A, C) because they model the appearance and the dynamics of the system. The matrix B, on the other hand, models the input noise and is not that critical to describe the appearance of the scene. Thus, a suitable candidate for the canonical form is the Observability Canonical Form (OCF) [16] given above by However, the estimation of the transformation that converts a set of parameters to this canonical form is numerically unstable [16]. As a result, in the presence of noise, two dynamical systems that are similar can be mapped to dynamical systems in the canonical form that are fairly different. In order to address this drawback, we propose to use a canonical form based on the Jordan real form. When A has $2q$ complex eigenvalues and $n-2q$ real eigenvalues, the Jordan Canonical Form is given by

$$A_c = \begin{bmatrix} -a_{n-1} & & & & \\ & -a_{n-2} & & & \\ & & \ddots & & \\ & & & -a_1 & \\ & & & & 0 \end{bmatrix}, \quad C_c = \begin{bmatrix} -a_0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{1 \times n} \quad (11)$$

$$A_c = \begin{bmatrix} \sigma_1 & \omega_1 & 0 & \dots & 0 & -\omega_1 & \sigma_1 & 0 & \dots & 0 & \vdots & \vdots & \vdots & 0 & 0 & 0 & 0 & \lambda_{2q-n-1} & 0 & 0 & 0 & \dots & 0 & \lambda_{2q-n} \\ & \end{bmatrix}, \quad C_c = [1 \ 0 \ 1 \ 0 \ \dots \ 1 \ 1], \quad (12)$$

where the eigenvalues of A are given by $\{\delta_1 \pm i\omega_1, \delta_2 \pm i\omega_2, \dots, \delta_q \pm i\omega_q, \lambda_1, \dots, \lambda_{n-2q}\}$. It can be noted that the JCF is indeed equivalent to the RCF or the OCF, but in a different basis.

Given any general canonical form based on A and C, we now outline the steps to convert the identified parameters into the canonical form. Assume that we have the identified parameters (A, C) . We now need to find an invertible matrix P such that $(PAP^{-1}, P^T C P^{-1}) = (A_c, C_c)$, where the subscript c represents any canonical form. The vector $\gamma \in \mathbb{R}^p$ is an arbitrary vector chosen to convert the LDS (A, C) with p outputs to a canonical form, which is defined for only one output. In our experiments, we set $\gamma = [1 \ 1 \dots 1]^T$ so that all rows of C are weighted equally. The relation between the A matrix and its canonical form A_c is a special form of the Sylvester equation:

$$A_c P - P A = 0. \quad (13)$$

Vectorizing this equation, we can solve for P as

$$\text{vec}(P) = \text{null} (I \otimes A_c - A \otimes I), \quad (14)$$

Where \otimes represents the Kronecker product. Similarly, if we consider the equation between the C matrices, $C^c P = \gamma^T C$, and vectorize it, we can solve for P by concatenating the two sets of equations as follows:

$$[I \otimes A_c - A \otimes I \ \ I \otimes C C] \text{vec}(P) = [0 \ C] \quad (15)$$

Once we have solved this equation, we can convert the parameters into the canonical form using P. It should be noted that the JCF is unique only up to a permutation of the eigenvalues. However, if we select a predefined order to sort the eigenvalues, we obtain a unique JCF.

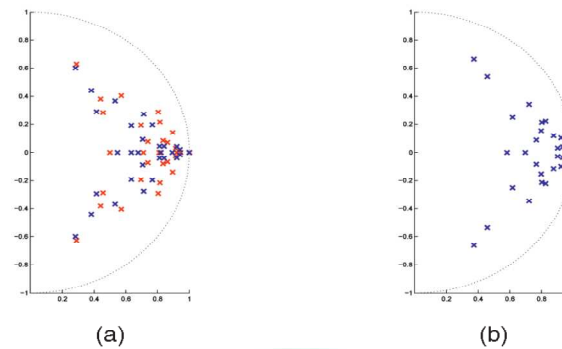
JOINT IDENTIFICATION OF DYNAMIC TEXTURES

In the prior section, we have shown how to convert the identified parameters into the same basis so that we can compare the C matrices to recover the registration. However, comparing the C matrices to recover the spatial transformation is based on using the assumption that the A matrices for the two systems were the same. This assumption is valid as we observe the same scene, and since the hidden states $z(t)$ captures the scene dynamics, they must evolve in the same way, irrespective of the viewpoint.

However, if we identify the LDS from each video sequence separately due to the presence of noise, viewpoint changes, and the suboptimal identification, there is no guarantee that the A matrix for the video sequences will be the same. This can be seen in Fig. 4a. We see that the eigenvalues of the A matrices identified separately are not the same. However, they are close to each other.

In this section, we propose a simple method to explicitly enforce the dynamics of multiple LDSs to be the same.

FIGURE 4



Eigenvalues of the identified A matrices of two spatially and temporally transformed video sequences. The red crosses denote the eigenvalues for sequence 1 and the blue crosses denote the eigenvalues of sequence 2. (a) Separate identification. (b) Joint identification.

Consider M video sequences, each represented as $I_i(t) \in \mathbb{R}^{p_i}, t \in \{1 \dots F\}, i \in \{1 \dots M\}$. Let us introduce $\tilde{I}_i(t) = I_i(t) - C_i^0$ for notational brevity. The traditional identification works by first forming the matrix $W_i = [\tilde{I}_i(1) \dots \tilde{I}_i(F)]$, and then, calculating the singular value decomposition of $W_i = U_i S_i V_i^T$. The parameters of the LDS are identified as $C_i = U_i(:, 1:n)$ and $Z_i = S_i(1:n, 1:n) V_i(:, 1:n)^T$. In our approach, we instead stack all the videos to form a single W matrix and factorize it using the SVD as

$$W = \begin{bmatrix} \tilde{I}_1 \\ \vdots \\ \tilde{I}_M \end{bmatrix} = \begin{bmatrix} U_1 \\ \vdots \\ U_M \end{bmatrix} \begin{bmatrix} S_1 \\ \vdots \\ S_M \end{bmatrix} \begin{bmatrix} V_1^T \\ \vdots \\ V_M^T \end{bmatrix} = USV^T \quad (16)$$

Although this seems to be the intuitively obvious thing to do, we will now show that this is indeed the correct thing to do. If, for the sake of analysis, we ignore the noise terms, we obtain the state evolution as $Z(t) = A^t Z_0$, where Z_0 is the initial state of the system. Now if we consider the temporal lag $T \in \mathbb{Z}$ for the i th video sequences, then the evolution of the hidden state of the i th sequence is given by $z_i(t) = A^T z_i(t)$. Therefore, we can now decompose W using the SVD as follows:

$$\begin{aligned} W &= C_1 A^T Z_1 \dots C_1 A^T Z_F \\ &= C_1 A^T : C_1 A^T Z_F \\ &= C_1 A^T : C_1 A^T Z_F \end{aligned} \quad \begin{aligned} &C_1 A^T Z_1 \dots C_1 A^T Z_F \\ &C_1 A^T Z_1 \dots C_1 A^T Z_F \end{aligned} \quad (17)$$

From the above equation, we can estimate a single common state for all the sequences. Moreover, given Z , we estimate a common dynamics matrix for all the sequences. Now we can also recover C_i from C up to the matrix A^T . The problem is that T_1 is unknown, so we cannot directly compute C_i from C . Now if we consider the equation for the i th video sequence, we can see that

$$[\tilde{I}_i(1) \dots \tilde{I}_i(F)] = C_i A^T [Z(1) \dots Z(F)] \quad (20)$$

$$= C_i A^T (A^T)^{-1} [Z(T_1+1) \dots Z(F+T_1)] \quad (21)$$

Thus, we see that the parameters we estimate are the original parameters of the system, but in a different basis. Therefore, by converting the parameters to the canonical form, we can remove the trailing A^T and recover the original parameters in their canonical form. The joint identification algorithm is outlined in Algorithm 1. Now, by construction, the A matrices of the multiple video sequences are the same. Hence, their eigenvalues are also the same. This can be seen in Fig. 4b.

Algorithm 1. Joint identification of video sequences

1 Given m video sequence $\{I_i(t) \in \mathbb{R}^{p_i}\}_{i=1}^m$, calculate the temporal mean of each sequence $C_i^0 \in \mathbb{R}^{p_i \times 1}, i=1, \dots, m$

2 Compute $C; Z$ using the rank n singular value decomposition of the matrix

$$W = \begin{bmatrix} \tilde{I}_1 \\ \vdots \\ \tilde{I}_m \end{bmatrix} = \begin{bmatrix} U_1 \\ \vdots \\ U_m \end{bmatrix} \begin{bmatrix} S_1 \\ \vdots \\ S_m \end{bmatrix} \begin{bmatrix} V_1^T \\ \vdots \\ V_m^T \end{bmatrix} = USV^T \quad (18)$$

$$Z = SV^T, C = U \quad (19)$$

3 Compute

$$A = [z(2), \dots, z(F)][z(1), \dots, z(F-1)]^T \in \mathbb{R}^{n \times n}$$

4 Let $C_i \in \mathbb{R}^{p_i \times n}$ be the matrix formed by rows $\sum_{j=1}^{i-1} p_j + 1$ to $\sum_{j=1}^i p_j$ of C , and convert the pair (A, C_i) to Jordan canonical form.

Having identified a dynamic texture model for all video sequences with a common A and all C matrices with respect to the same basis, in the next section we describe a method to register multiple video sequences using the appearance parameters of the LDSs. Other applications of using the joint identification include recognition of dynamic textures, joint synthesis of videos, etc.

REGISTERING USING THE DYNAMIC TEXTURE MODEL

In this section, we propose an algorithm to recover the spatial transformation from the appearance parameters of the LDSs identified from the two video sequences. We need to compare this parameter between two LDSs, to recover the relative spatial alignment. In our paper, we term the mean image (C_0) and the n columns of the C matrix as the dynamic appearance images. Let us consider two video sequences, $I_1(x, t)$ and $I_2(x, t)$, where x denotes the pixel coordinates and $t=1, \dots, F$. We assume that the video sequences are related by a Homography H and a temporal lag T , i.e., $I_1(x, t) = I_2(H(x), t+T)$. Once we recover the spatial alignment independent of the temporal lag between the video sequences, we temporally align the two sequences using a simple line search in the temporal direction, i.e., $T = \arg \min_t \sum_i \|I_1(x, t) - I_2(H(x), t+T)\|^2, T \in \mathbb{Z}$.

Our algorithm to spatially register the two video sequences $I_1(t)$ and $I_2(t)$ proceeds as follows: We calculate the mean images C_1^0 and C_2^0 , identify the system parameters (A, C_1) and (A, C_2) in the JCF, and convert every column of C_i into its image form. We use the notation C_i^j to denote the i th column of the j th sequence represented as an image. We use a feature-based approach to spatially register the two sets of images $\{C_1^0, C_1^1, \dots, C_1^n\}$ and $\{C_2^0, C_2^1, \dots, C_2^n\}$. We extract SIFT features and feature descriptor around every feature point in the two sets of $n+1$ images. We match the features extracted from image C_1^j with those extracted from image C_2^j , where $j \in \{0, \dots, n\}$, i.e., the forward direction. We also match the features from C_2^j with those extracted from image C_1^j , where $j \in \{0, \dots, n\}$, i.e., the reverse direction. We retain only the matches that are consistent both in the forward direction and the reverse direction. We then concatenate the correspondences into the matrices $X_1 \in \mathbb{R}^{3 \times M}$ and $X_2 \in \mathbb{R}^{3 \times M}$. The corresponding columns of X_1 and X_2 are the location of the matched features in homogenous coordinates and M is the total number of matches from the $n+1$ image pairs. We then need to recover a homography H such that $X_2 \sim HX_1$. In order to recover the homography, we first run RANSAC and obtain the inliers from the matches. We then fit a homography using the nonlinear method outlined in [17]. Our registration algorithm is summarized in Algorithm 2.

Algorithm 2. Registration of video sequences

1 Given $I_1(t)$ and $I_2(t)$, calculate the parameters A, C_1^0 , and C_2^0 .

2 Extract features and the descriptors from $(C_i^j), j = \{1, 2\}, i = 0, \dots, n$.

3 Match features from C_1^j to C_2^j and also in the reverse direction. Retain the matches that are consistent across both directions and concatenate the feature point location from C_1^j into X_1 and its corresponding match into X_2 .

4 Recover the homography H using RANSAC such that $X_2 \sim HX_1$.

5 Calculate temporal alignment T as $T = \arg \min_t \sum_i \|I_1(x, t) - I_2(H(x), t+T)\|^2$.

Remark For rigid sequences, one could potentially argue that the mean image would be sufficient to register the video sequences. This motivates the fact that for such sequences, we could use only the matches from the mean images rather than the matches from C matrix when estimating the registration parameters.

The scenario for nonrigid scenes is the exact opposite. We would like to use the matches from C matrix rather than the matches from the mean image. Note that in our algorithm, the best matches given by RANSAC could arise from the mean image or the dynamic appearance images or both. Hence, we do not explicitly restrict the algorithm to use only the mean image or only the dynamic appearance images, as done in [15]. This choice now becomes automatic and makes our method applicable to both rigid and nonrigid sequences.

EXPERIMENTAL RESULTS

OBJECT DETECTION

For this section, the algorithm developed in this paper, which builds object information as a consequence of the registration, is compared against a difference-based object detection method that is applied post-registration.

Figure 3 shows a pair of images that will be used for this test. The images have a moving object, and the camera is shifted and rotated between frames. Block size of 20×20 is used.

Upon application of the first object detection method, the result is shown in Figure 6. The image displayed is the mean of the images once they have been registered and aligned, allowing the two locations of the object to be visible simultaneously. The blocks identified as not belonging to the background give a fair approximation of the object. Figure 5 shows the same sequence processed by the difference-based method. The results are not massively different, with the registration-based detection giving slightly better coverage of the object. The main difference is that the difference-based method required a significant amount of threshold tweaking to get to this point, where the registration method did not.

BLOCK SIZE

The block sizes tested for the current incarnation of the algorithm is all square, mainly to keep things simple. The use of rectangular and other shaped blocks at this stage is unlikely to have any benefit.

The block sizes looked at for the test sequences (which measure 320×240 pixels) are as follows: 5×5 , 10×10 , 15×15 , 20×20 , 25×25 , 30×30 , and 40×40 . The same sized block is applied to the reduced-scale images used for the initial coarse search conducted by the full-spread search algorithm (see Section II-B), as the minimum number of pixels required to produce accurate results does not change as the image gets smaller, since the size of a pixel remains constant. Figure 6 shows the mean sum of squares error for registration with different block sizes over three different image pairs.

When the algorithm is run with block sizes of 5×5 and 10×10 , the number of pixels in each block is insufficient to accurately distinguish unique sections of the image and correctly map them in the registration. As a result, the registration at these block sizes is not sufficiently accurate to be useful. The object detection at this level is also not very effective, as essentially anywhere in the image with significant edges are picked up as deviating from the overall transform (since the transform is incorrect).

FIGURE 5



Sample image pair for registration involving camera motion and an independently moving object.

FIGURE 6: OBJECT DETECTION USING RESULTS OF THE BLOCK-BASED REGISTRATION



FIGURE 7: OBJECT DETECTION USING DIFFERENCE-BASED METHOD

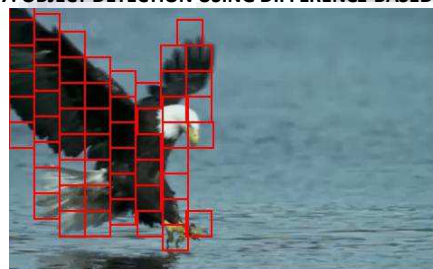
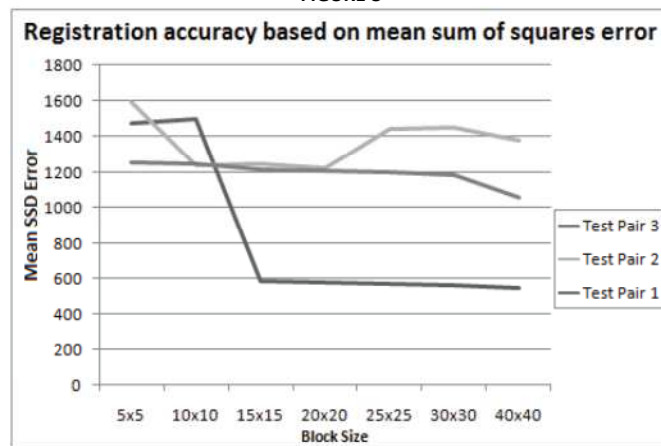


FIGURE 8



This shows the accuracy of the registrations on different test images using different block sizes. Three image pairs were used, based on images with different image statistics. Error values are mean sum-of-squares difference.

With a block-size of 15×15 to 25×25 the registration result remains the same for most image sequences, but in some cases 20 × 20 will give a more accurate result. In this case it can become a trade-off situation, since the smaller block size will give better object detail, but in most cases registration accuracy is the crucial element so a block size of 20 × 20 is best.

Beyond a block size of 30×30, the graph indicates that the registration is effective; however this is a limitation of using sum-of-squares difference as a measure of error when there are multiple objects in the image – the actual registration result for two of the test pairs is very poor, but the error value is low. The problem with a larger block size is that the proportion of blocks that contain multiple objects is high, which interferes with the registration.

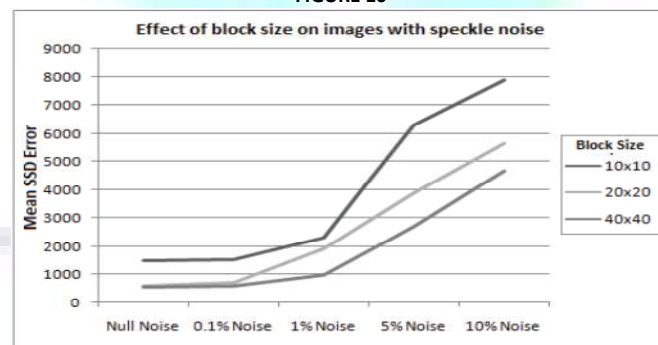
Another consideration for the algorithm is the capability for dealing with noise. This capability is dependent on the type of noise present: if the noise is in the form of a small number of clustered noise pixels across the image, these clusters will simply be treated as separate objects and captured by the algorithm; with speckle noise (see Figure 7), however, the entire registration can be affected. Figure 8 shows the accuracy of the registration for images containing different levels of speckle noise using different block sizes. This indicates that larger block sizes enable the algorithm to more easily cope with high levels of noise.

FIGURE 9



Example image with 5% speckle noise. The amount of damage is substantial.

FIGURE 10



Different amounts of speckle noise applied to sample images. This graph shows the accuracy of the registration using different block sizes for different noise levels. The difference in accuracy between the block sizes does not change very much over the changing noise levels, and beyond about 1% noise the accuracy is degraded too much to be useful.

CONCLUSION

Dynamic texture of registering video sequences for recovering the spatial transformation independent of temporal transformation. Our result in the research shows that the methods reduce that number of frames at a time in one sequence. One needs to perform future extraction tracking and trajectory matching the multiset of F frames. In our case, we only need some extractions over for multi sets $(n+1) \times \dots \times n$ images. In addition we have computation for the system parameters typically for F/n for the sequences presented in the range of 10-15 therefore using our method gives an advantage in with respect to number of frames are merged at the time. In short, we have presented a method that works equally well compare to the state of the art but is more efficient.

The other two important contributions of this paper are the use of a joint together of system identification framework with a canonical form representation. The joint identification and the Jordan canonical form are not only applicable to the case of registering video sequences, but also to the entire genre of algorithms based on the dynamic texture model. In this paper, we have also shown that out of all the possible choices for the method of identification and canonical form, the JID using JCF performs the best.

REFERENCES

1. B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, October 2003.
2. Bin Qi, M. Ghazal, and A. Amer. Robust global motion estimation oriented to video object segmentation. *Image Processing, IEEE Transactions on*, 17(6):958–967, June 2008.
3. Chan and N. Vasconcelos, “Probabilistic Kernels for the Classification of Auto-Regressive Visual Processes,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 846–851, 2005.
4. Fang Zhu, Ping Xue, and Eeping Ong. Low-complexity global motion estimation based on content analysis. In *Circuits and Systems, 2003. ISCAS '03. Proceedings of the 2003 International Symposium on*, volume 2, pages II–624–II–627 vol.2, May 2003.
5. G. Doretto, A. Chiuso, Y. Wu, and S. Soatto, “Dynamic Textures,” *Int’l J. Computer Vision*, vol. 51, no. 2, pp. 91–109, 2003.
6. John Y. A. Wang and Edward H. Adelson. Representing moving images with layers. *IEEE Transactions on Image Processing*, 3:625–638, 1994.
7. Jose Miguel Buenaposada, Enrique Munoz, and Luis Baumela. Tracking a planar patch by additive image registration. 2003.
8. Krutz, M. Frater, M. Kunter, and T. Sikora. Windowed image registration for robust mosaicing of scenes with large background occlusions. In *Image Processing, 2006 IEEE International Conference on*, pages 353–356, Oct. 2006.
9. Krutz, M. Frater, and T. Sikora. Window-based image registration using variable window sizes. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 5, pages V–369–V–372, 16 2007–Oct. 19 2007.
10. M. Shah and R. Kumar (Eds). *Video Registration*. Kluwer, May 2003.
11. Marina Georgia Arvanitidou, Alexander Glantz, Andreas Krutz, Thomas Sikora, Marta Mrak, and Ahmet Kondo. Global motion estimation using variable block sizes and its application to object segmentation. *Image Analysis for Multimedia Interactive Services, International Workshop on*, pages 173–176, 2009.
12. P. Zhilkin and M. E. Alexander. A patch algorithm for fast registration of distortions. *Vibrational Spectroscopy*, 28(1):67 – 72, 2002.
13. P.V. Overschee and B.D. Moor, “Subspace Algorithms for the Stochastic Identification Problem,” *Automatica*, vol. 29, no. 3, pp. 649–660, 1993.
14. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge Univ. Press, 2000.
15. R. Vidal and A. Ravichandran, “Optical Flow Estimation and Segmentation of Multiple Moving Dynamic Textures,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 516–521, 2005.
16. Ravichandran and R. Vidal, “Mosaicing Nonrigid Dynamical Scenes,” *Proc. Workshop Dynamic Vision*, 2007.
17. W.J. Rugh, *Linear System Theory*, second ed. Prentice Hall, 1996.

REQUEST FOR FEEDBACK

Dear Readers

At the very outset, International Journal of Research in Commerce, IT and Management (IJRCM) acknowledges & appreciates your efforts in showing interest in our present issue under your kind perusal.

I would like to request you to supply your critical comments and suggestions about the material published in this issue as well as on the journal as a whole, on our E-mails i.e. **infoijrcm@gmail.com** or **info@ijrcm.org.in** for further improvements in the interest of research.

If you have any queries please feel free to contact us on our E-mail infoijrcm@gmail.com.

I am sure that your feedback and deliberations would make future issues better – a result of our joint effort.

Looking forward an appropriate consideration.

With sincere regards

Thanking you profoundly

Academically yours

Sd/-

Co-ordinator

ABOUT THE JOURNAL

In this age of Commerce, Economics, Computer, I.T. & Management and cut throat competition, a group of intellectuals felt the need to have some platform, where young and budding managers and academicians could express their views and discuss the problems among their peers. This journal was conceived with this noble intention in view. This journal has been introduced to give an opportunity for expressing refined and innovative ideas in this field. It is our humble endeavour to provide a springboard to the upcoming specialists and give a chance to know about the latest in the sphere of research and knowledge. We have taken a small step and we hope that with the active co-operation of like-minded scholars, we shall be able to serve the society with our humble efforts.

Our Other Journals

