A Review of the Advances in Population Estimation and Human Behaviour Estimation Using Computer Vision and Mobile Phone Data

Anurag Unnikrishnan[#1], Anshuman Narayan[#2], HS Guruprasad [*3]

[#]UG Scholar, Department of CSE, BMSCE

*Professor and Head, Department of CSE, BMSCE
Bangalore, India

Abstract — *Owing to the population boom in the past decade, the need for efficient and accurate crowd control techniques have risen by the dozen. To prevent violent riots and stampedes, these techniques can provide useful insight into crowd density and location and can help the authorities in channelling their resources efficiently. This paper is a survey of all the various techniques that incorporate computer vision and mobile traffic data analysis. An understanding of these is vital in establishing a formidable crowd control model.*

Keywords — *computer vision, GLCM, textural classifier, GPS.*

I. INTRODUCTION

Leading the charge at the dawn of a new technological revolution are computer vision and artificial intelligence, among others. This has given rise to a shift in dynamics, from human-centred applications to stand-alone artificial agents that can bring about a performance that is far more variant in execution.

Computer vision is at the heart of a plethora of applications, right from face recognition to inspecting bottles on a product line. It is constantly used by autonomous robotic systems to identify their surroundings and enable them to manoeuvre their way through any environment. However, the main problem that arises when dealing with systems that rely on computer vision is the ability to discern a scene. Understanding the image and performing the required analysis are the critical aspects of computer vision.

Gone are those days when human presence was required to perform manual tests on a basic prototype, control the robotic milieu in a factory or the like. Artificial intelligence has made human lives easier by automating the daily tasks. Though still a niche concept, artificial intelligence and computer vision go hand in hand. Riding on the coattails of this union, this paper covers the techniques used by both these concepts to estimate the human population of a given region. The applications of these techniques lie in observing the population at large gatherings or religious institutions to ensure public safety by allowing regulatory authorities to take precautionary measures on observed instances of increasing population density,

This paper is a literature survey of the various concepts that are needed to successfully track people using computer vision. It also discusses the methods used to analyse mobile traffic data to discern the population density in a given area at a given time. It also discusses some of the techniques that have been proposed by experts in the field and expands on the possibility of using both techniques simultaneously.

II. TECHNIQUES INVOLVING COMPUTER VISION

While there exist many techniques in estimating the size of a population using computer vision, Davies et al [1] discuss the application of general image processing to identify the density of a given population in each area.

One novel approach chosen by many authors, is the use of a textural classifier to solve the problem. Liu and Clarke [2] discuss the usefulness of a textural classifier that is based on a Gray Level Co-occurrence Matrix (GLCM). Using the classifier, detailed land use information is extracted. The conclusion is that texture alone, in remote image sensing is not sufficient in obtaining feasible results and it needs to be supplemented with other information, like, spectral reflectance, to achieve high accuracy.

Yang et al [3] propose using Support Vector Machines(SVM) to arrive at an estimate of the crowd density along with linear regression to allow for more accuracy. Gupta and Gupta [4] also propose a similar model in their work. While GCLMs are a reliable technique, Wang et al [5] propose an alternative that improves on the performance issues of GLCM and makes it more suitable for real-time use. They propose a method called Principal Component Analysis (PCA) to obtain the input vectors and infer the population from these inputs. Wen et al [6] propose another alternative to GCLMs to overcome the problem of overlapping of people in crowds and any occlusion, using Gabor filters to obtain texture features. These features are used in conjunction with the count of the number of people in an area to arrive at an empirical relationship. This relationship is learned through SVMs to arrive at an accurate approximate to the number of people in a crowd.

Crowd Analysis is a well-researched domain that aims to tackle the issue of controlling large crowds and identifying abnormal behaviours. Sjarif et al [7] discuss the process of inspecting abnormal behaviour in high density crowded areas by overcoming obstacles such as changes of motion, background differences, irregularities in illumination etc. by exploiting the advances in computer vision, pattern recognition and data mining. Andrade and

Fisher [8] discuss in their paper the approaches that are capable of simulating and hence providing visual evidence of threatening crowd related calamities such as overcrowding due to a bottle-necked exit, stampedes, panic induced disturbances etc. These simulations can then be used to verify and validate crowd control techniques. Without such simulations, reproducing those situations in reality is not only risky but also unsafe. Gupta and Gupta [9] also discuss how occlusion can be overcome by using, continuous monitoring and evaluation of recurring patterns.

Caesar and Musse [10] provide an array of approaches to tackle the recurring problems that arise in crowd analysis such as people density evaluation, tracking people in crowded areas and understanding the behaviour of these crowds. They show how an object-level model can provide an accurate understanding of the count in a geographic location. Joshi and Vohra [11] incorporate the fields of sociology and psychology to create similar models. Sirmacek and Reinartz [12] discuss a model that uses a probability framework that is feature dependent. Redundant features are eliminated from the input image and then the relevant features are used to estimate the crowd density using kernel based regression.

While crowd analysis is a promising field that provides great insight into using computer vision to analyse and quantify large crowds, movements within such a collective, act as a hindrance to the analysis. Gavrila [13] talks about analysis of various 2D and 3D approaches to analyse human movement. While 2D model approaches may differ in the use of shape models, 3D models are simpler to analyse as priori knowledge of kinematics and structural properties of human bodies reduce the problem to a classification one. Zhang and Li [14] propose a method to overcome random motions of the public in crowded areas by using an Accumulated Mosaic Image Difference(AMID) to arrive at an algorithm for segmenting the foreground. This foreground representation is then analysed with a deviation model to arrive at a number for the crowd density.

Stauffer and Grimson [15] provide an alternative to motion tracking by learning the patterns in the motion of the object. To segment the motion, an adaptive background subtraction is used. This method mimics each pixel as separate Gaussians distribution and uses this approximation to update a model. An evaluation of these Gaussian distributions is conducted to estimate the most likely model. This results in a real-time and stable tracker that is immune to changes in lighting, clutter motions that are relative and erratic changes in scenes.

McKenna et al [16] discuss a method of using a background subtraction algorithm along with the combination of color and information of gradients to deal with shadows and irregular color prompts. Successful tracking of multiple people requires specific models that can deal with occlusions to a certain extent. More emphasis is given to a color-based model because color and texture do propose the decomposition of an individual into various visual parts. Also, the distribution of color on any item of clothing on a person is always stable irrespective of their orientation and is more resilient to changes in illumination.

III. Techniques Involving Mobile Phone Data

An alternative method to estimating human population involves the counting of mobile phone nodes in each region and using mathematical formulas and functions to approximate the human population in the region. It is a feasible approach considering the high mobile phone penetration rates across the globe.

GPS or Global Positioning System is a satellite based navigation system that can provide real-time location of any GPS receiver. Tsubouchi et al [17] propose an approach for calculating the population of a given area using GPS data from willing users. The algorithm uses data elimination steps that interpret data to exclude users inside buildings and commuters on vehicles. It also uses Gaussian kernels to counter for spatial and temporal scarcity of data allowing for the system to predict population figures at non-peak hours and in remote and larger areas.

Herrera et al [18] propose an alternative to GPS based estimation by using Virtual Trip Lines (VTLs). VTLs are virtual lines draw on points of interest which trigger sending GPS information to the server. The VTLs also use static information among themselves to produce more realistic and real-time traffic data. The VTL data obtained from the experiment successfully calculated and estimated car population and average velocity along a strip of a busy intersection and provided accurate information during peak traffic hours.

While GPS can prove to be an extremely valuable tool to estimate human population, it is a technology found in mostly smartphones that immediately alienates most of the mobile phone user group who may use cheaper handsets. It is therefore prudent to consider using records that can be generated by all mobile phones without the need for external applications. Call Detail Record(CDR) is a record generated by telecommunication service providers for each call made or text message sent.

These records are inherently anonymous as they contain only the mobile phone number and the cell phone tower used while making the call. Raslan and Elragal [19] propose a method of calculating population distribution of a region with the help of CDRs. This data is used in conjunction with knowledge of the region in question to arrive at POI (Points of Interest). These POIs are then used as markers around which population distributions and patterns can then be predicted. Khodabandelou et al [20] describe a similar model that uses cell phone tower generated metadata of a region to estimate its population. The model employs a variety of data filters such as population distribution, land use estimation using MWS, and static population data filters such as hour filtering and day filtering to achieve favourable accuracy in predicting the population of a given region. The authors test this model against multiple scenarios and achieve favourable results.

Deville et al [21] in their work, employ a similar model, using network tower based mobile information along with corrections for spatial and temporal distributions of this data. The results showed that the proposed method of population mapping was more accurate than traditional techniques using simple adjustments for spatial and temporal variations in data. The approach could also be modified to regions with lower mobile phone penetration with the help of added constants and formula to estimate the population using subprime mobile phone data.

CDRs provide relevant and clearer insight into the estimation of population of regions that suffer from economic adversity. Smith-Clarke et al [22] discuss the use of CDRs to estimate the socio-economic conditions of the surveyed region. The authors propose a model that uses the CDR database and feature extraction techniques such as gravity models, network advantage and introversion to estimate the population density and their activity and thus arrives at conclusions on the region’s socio-economic condition.

Use of mobile tower allows for finer distinction of poor or more affluent communities. Wesolowski and Eagle [23] describe their study on using data gained from mobile phone usage to analyse the population flow and behavioural patterns of human beings in slums. They conducted a case study using data from the Kibera in Nairobi, Kenya and were able to extract meaningful information about the population residing in the slums. Performing the proposed analysis of the movement of the population allowed for identification of social information such as tribal affiliations.

The rapid availability of CDRs finds its use in disaster and disease management. Bengtsson et al [24] discuss their work in creating a model that can extract location information of large population using the information gained from the SIM (subscriber Identity Module) cards wherein each SIM card’s location is tracked using CDR that pinpoint the nearest cell phone tower. This model was deployed to track movements of the populace during and after the Haiti earthquake. The accuracy of the model was tested against local government agency records and it was concluded that the model provided rapid and accurate population information that is critical for relief measures.

Wilson et al [25] elaborate on the successful use of mobile phone data i.e. CDR to accurately monitor the population and its movement in Nepal when it was affected by earthquakes during 2015. They describe the pre-processing of raw data and expand on data filters used to accurately discern population flows that provides valuable information to various humanitarian agencies. CDRs can also be used to study the spreading of commutable diseases and thereby help in tackling an epidemic.

Frias-Martinez et al [26] propose a model to predict the spreading of commutable diseases using data from individual mobile use patterns and social networks extracted from cell phone records. The model uses an agent based epidemic model to simulate a virus on a simulated environment created by the data generated from records. The model was simulated using data from the 2009 H1N1 flu outbreak in Mexico and the impact of mobility restriction on the spreading of the virus was analysed to conclude that mobility restriction greatly reduced the spreading of the disease.

CDRs can often be used in conjunctions with other datasets to support a given model’s estimate. Douglass et al [27] propose a method for population estimation of a given region with the help of telecommunication data in conjunction with census data and satellite images. They develop a model that formulates a relationship between the density of population of a region against the call volumes and use information gained from state conducted census and satellite images to arrive at a satisfactory estimate that can accurately predict population patterns and density. An experimental study on predicting the population density of the city of Milan in Italy is done using these techniques and the results speak favourably of the model’s accuracy.

CDRs also find their use in understand human and traffic movement and analysing human settlements. Berlingerio et al [28] discuss a model for using mobile phone data of the population of a specified region to compute the population and traffic density across the region and estimate each individual’s origin and destination flow and use these patterns to optimize public transportation and traffic in the region by providing users with alternative delivery routes that ensure that traffic is managed well.

Work and Bayen [29] discuss a similar model that uses cell phone GPS data instead.
CDRs used in conjunction with statistical records and statistical analysis allow for a clearer picture to account for aberrations in CDR data. Furletti et al [30] propose a model that is an analysis process on top of an existing project called Sociometer, a data mining tool that classifies people on the data extracted from their call records. While this tool is restricted to one area, the authors' modifications allow for the tool to be used for scanning multiple areas and calculating work flows between these areas. The analysis process involves creating profiles of users based on usage patterns and using these patterns to identify users who are residents of a region versus commuters. This classification provides significant insight into the behavioural patterns of the studied population. The model was tested in the province of Pisa in Italy and results of the analysis were compared to another project that used administrative data sources.

De Jonge et al [31] elaborate on a model that can calculate the live dynamic populations of regions using the data generated from mobile phone usage and using smart data filters and applying data processing techniques such as building cells as a basic unit in the model and inferring information about the users themselves by analysing their mobile usage patterns. The analysis also allows for classifying groups of cells as residential or commercial blocks.

Data records of government administrative agencies often act as a tool to further solidify estimates that are drawn from live sources such as CDRs. McPherson and Brown [32] discuss various population estimation measures and compare them. They propose a method that involves extrapolation of data from publicly available census data and creation of a population density map of U.S Cities using location data publicly available for each County. The proposal involves dividing the maps into grids by performing calculations on census data and assigning values to these grids to reflect the densities. The results were compared against the county data to check if any errors had been added by the computations. The results showed that the model was successful in performing extremely well in the calculation of workplace and night-time population. The model showed correlation coefficients for the two populations are 0.99 and 1.6. The authors also concluded with discussing the uncertainty of using publicly available data due to it being prone to errors.

IV. Conclusions

It can therefore be concluded that it is indeed possible to create such a model by manipulating the concepts of computer vision and supplementing the resulting analysis with the data recovered from mobile towers. Aided with artificial intelligence, such a model can be deployed on a large scale.

While using GPS to track the location allows for more accuracy, the security issues and uncertainty in availability of GPS enabled phones alludes that CDRs are a more viable alternative despite a decrease in accuracy. The high density of cell phone towers and the success of using data filters to offset the shortcomings in accuracy concludes that CDRs reflect an accurate estimate of the population.

References

[1] Davies C.A, Yin J.H and Velastin S.A," Crowd Monitoring Using Image Processing", IEEE Electronic Communications Engineering Journal, Vol 7, No. 1(Feb), pp 37-47.

[2]] Liu, X., and K.C. Clarke, 2002. "Estimation of residential population using high resolution satellite imagery," Proceedings of the 3rd Symposium on Remote Sensing of Urban Areas (D. Maktav, C. Juergens, and F. Sunar-Erbek, editors), 11–13 June, Istanbul, Turkey (Istanbul Technical University), pp. 153–160.

[3] Yang J, Jin L and He Y," Crowd Density and Counting Estimation Based on Image Textural Feature", Journal of Multimedia, Vol 9, No. 10, October 2014.

[4] Gupta J.K and Gupta S.K," Design and Analysis of Crowd Sized Estimation Techniques", International Journal of Computer Applications, Vol.107, No. 19, December 2014.

[5] Wang B, Bao H, Yang S and Lou H, "Crowd Density Estimation Based on Texture Feature Extraction", Journal of Multimedia, Vol. 8, No. 4, August 2013.

[6] Qing Wen, Chengcheng Jia, Yangquan Yu, Gang Chen, Zhezhou Yu and Chunguang Zhou," People Number Estimation in the Crowded Scenes Using Texture Analysis Based on Gabor Filter", Journal of Computational Information Systems 7: 11 (2011) 3754-3763.

[7] Sjarif N.A.A, Shamsuddin S.M, Hashim S.Z," Detection of Abnormal Behaviours In Crowd Scene: A Review", International Journal of Advances in Soft Computing and its Applications, January 2012.

[8] Andrade E, Fisher R, "Simulation of Crowd Problems for Computer Vision, First International Workshop on Crowd Simulation, vol. 3, pp. 71-80 (2005).

[9] Gupta J and Gupta S.K," Approach Based Study of Crowd Size Analysis", International Journal of Research in Advent Technology, Vol 2, No. 3, March 2014 E-ISSN: 2321-9637.

[10] Caesar J and Musse S," Crowd Analysis Using Computer Vision Techniques", IEEE Signal Processing Magazine, Volume 27, Issue 5, Sept 2010.

[11] Joshi K.Y and Vohra S.A," Crowd Behaviour Analysis", International Journal of Science and Research, ISSN: 2319-7064.

[12] Sirmacek B and Reinartz P," Automatic Crowd Analysis from Very High Resolution Satellite Images",5th International Conferences on Recent Advances in Space Technologies,2011, ISBN:978-1-4244-9616-7.

[13] D. M. Gavrila. The visual analysis of human movement: A survey. CVIU, 73(1):82–98, 1999.

[14] Zhang Z and Li M, "Crowd density estimation based on statistical analysis of local intra-crowd motions for public area surveillance" Opt. Eng. 51(4), 047204 (May 02, 2012).

[15] Stauffer, C. and Grimson, W. 2000. Learning patterns of activity using real time tracking. IEEE Trans. Pattern Analysis Machine Intelligence 22, 8, 747–767.

[16] S. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler, "Tracking groups of people," Computer Vision Image Understanding, vol. 80, no. 1, pp. 42–56, 2000.

[17] Nishi K, Tsubouchi K, Shimosaka M. Hourly pedestrian population trends estimation using location data from smartphones dealing with temporal and spatial sparsity. In: Proc. of the 22nd ACM Int'l Conf. on Advances in Geographic Information Systems. Dallas/Fort Worth: SIGSPATIAL, 2014. 281−290. [doi: 10.1145/2666310.2666391].

[18] Herrera, J. C., Work, D. Ban, X., Herring, R. Jacobson, Q. and Bayen, A. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. Transportation Research C, 18, pp. 568--583, 2010.

[19] Raslan H and Elragal A," Inferring Urban Population Distribution from GSM Data: An Experimental Case Study", The Seventh International Conference on Advanced Geographic Information Systems, Applications, and Services, February 2015.

[20] Khodabandelou G, Gauthier V, El-Yacoubi A.M and Fiore M," Population Estimation from Mobile Network Traffic Metadata",17th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2016, DOI:10.1109/WoWMoM.2016.7523554.

[21] Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F.R., Gaughan, A.E., Blondel, V.D., Tatem, A. J.: Dynamic population mapping using mobile phone data. Proceedings of the National Academy of Sciences 111(45), 15888–15893 (2014).

[22] Smith-Clarke C, Mashhadi A and Capra L," Poverty on the cheap: estimating poverty maps using aggregated mobile communication networks", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Pages 511-520.

[23] Wesolowski A and Eagle N, "Inferring Human Dynamics in Slums Using Mobile Data", Technical Report, Santa Fe Institute, 2009.

[24] Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J (2011) Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti. PLoS Med 8(8): e1001083. doi: 10.1371/journal.pmed.1001083.

[25] Wilson R, zu Erbach Schoenberg E, Albert M, Power D, Tudge S, Gonzalez M, Guthrie S, Chamberlain H, Brooks C, Hughes C, Pitonakova L, Buckee C, Lu X, Wetter E, Tatem A, Bengtsson L. Rapid and Near Real-time Assessments of Population Displacement Using Mobile Phone Data Following Disasters: The 2015 Nepal Earthquake. PLOS Currents Disasters. 2016 Feb 24. Edition1.doi:10.1371/currents.dis.d073fbece328e4c39087bc086d694b5c.

[26] Frias-Martinez E, Williamson G and Frias-Martinez G," Agent-Based Modelling of Epidemic Spreading using Social Networks and Human Mobility Patterns." IEEE Third International Conference on Privacy, Security, Risk and Trust 2011, DOI:10.1109/PASSAT/SocialCom.2011.142.

[27] Douglass, R.W., Meyer, D.A., Ram, M., Rideout D., Song D.," High resolution population estimates from telecommunications data" EPJ Data Sci. (2015) 4: 4. doi:10.1140/epjds/s13688-015-0040-6.

[28] Berlingerio M, Calabrese F, Di Lorenzo G, Nar R, Pinelli F and Sbodio L.M," All Aboard: A System for Exploring Urban Mobility and Optimizing Public Transport Using Cellphone Data", Machine Learning and Knowledge Discovery in Databases, European Conference, ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part III, pp 663-666.

[29] B. Work, Daniel and M. Bayen Alexander, Impacts of the Mobile Internet on Transportation Cyberphysical Systems: Traffic Monitoring using Smartphones, National Workshop for Research on High-Confidence Transportation Cyber-Physical Systems: Automotive, Aviation and Rail Washington, DC, November 18-20, 2008.

[30] Furletti B, Gabbrielli M, Giannotti F, Milli L, Nanni M, Pedreschi D," Use of mobile phone data to estimate mobility flows. Measuring urban population and inter-city mobility using big data in an integrated approach", 47th SIS Scientific Meeting of the Italian Statistical Society Cagliari 2014, ISBN: 978-88-8467-874-4.

[31] de Jonge E, van Pelt M and Roos M," Time patterns, geospatial clustering and mobility statistics based on mobile phone network data", Federal Committee on Statistical Methodology Washington D.C, DOI:10.13140/2.1.1536.1920.

[32] McPherson, Timothy N. and Brown, M. Estimating Daytime and Nighttime Population Distributions in U.S Cities for Emergency Response Activities. Symposium on Planning, Nowcasting and Forecasting in the Urban Zone, January 2004.