

**KAJIAN SINGKAT TENTANG:
PENGENDALIAN DAN PENJEJAKAN OBJEK BERBASIS VISUAL**

***Ari Yulianti, Carmadi Machbub, Pranoto H. Rusmin**

Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung, Indonesia.

**Email: ari.yulianti@students.itb.ac.id*

Abstrak

Pengendalian dan penjejukan objek banyak digunakan di berbagai bidang seperti pendidikan, kesehatan dan kesejahteraan, olahraga, pengawasan pada industri konstruksi, pengawasan di supermarket, dan lain-lain. Sistem pengendalian dan penjejukan objek terdiri dari tiga sub-sistem, yaitu pendeteksi dan pengenalan objek, sistem estimasi pergerakan objek, dan pengendalian perangkat kamera dan aktuatornya agar dapat menjejaki objek. Banyak kajian telah dilakukan tentang beberapa algoritma untuk melakukan deteksi dan pengenalan objek seperti *camshift*, *viola-jones*, *gaussian mixture model*, *surf*, dan lain-lain. Dengan kelebihan dan kelemahan sebagai berikut (i) ketika data terbatas jumlahnya, (ii) sistem operasi berbasis fitur jauh lebih cepat daripada sistem berbasis piksel, (iii) posisi, ukuran, orientasi objek yang berubah-ubah, dan lain-lain. Algoritma-algoritma ini telah dikembangkan untuk estimasi pergerakan objek namun dalam hal ini masih terdapat kelemahan antara lain pada gerak objek dan lokasi objek yang selalu berubah-ubah dan jumlah objek lebih banyak daripada jumlah kamera. Sedangkan, persoalan yang dihadapi dalam sistem pengendalian adalah ketika penjejukan pada objek terjadi seperti (i) *hijacking problem*, (ii) *centralization problem*, (iii) *drifting problem*. Tujuan dari kajian ini adalah untuk melakukan bahasan dari penelitian mengenai pengendalian dan penjejukan objek, yang telah dilakukan sebelumnya dan untuk mendapatkan titik temu peluang kontribusi dan nilai kebaruan dari bidang ini.

Kata Kunci : pengendalian, penjejukan objek, estimasi

Abstract

The Control and object tracking is widely used in various fields such as education, health and welfare, sports, supervision on construction industry, supervision at supermarkets, etc. The object's control and tracking system consists of three sub-systems, namely object detection and recognition, estimation system of object movement, and control of camera devices and actuators till track objects. There have been research on some algorithms for object detection and recognition such as *camshift*, *viola-jones*, *gaussian mixture model*, *surf*, etc. By the following advantages and disadvantages as consequent (i) when data is limited in number, (ii) feature-based operating systems are much faster than pixel-based systems, (iii) position, size, object orientation is fickle, etc. These algorithms have been expanded to estimate the movement of objects still in this case there is just infirmity obiter on the movement of objects and the location of objects that are always changing and the number of objects more than the number of cameras. Whereas, the problem faced in the control system is when the tracking of the object occurs as though (i) *hijacking problem*, (ii) *centralization problem*, (iii) *drifting problem*. The purpose of this study is to conduct a discussion of the research on object control and tracking, which has been done previously and to derive a point of contribution chance and novelty value of this field.

Keywords : control, object tracking, estimation

1. Pendahuluan

Penjejakan objek meskipun ranahnya dalam teori kendali, telah semakin banyak yang tertarik terutama dalam berbagai aplikasi teknologi visi komputer selama dekade terakhir ini. Sistem penjejakan objek berbasis visual mampu menjejak sejumlah variabel objek dalam *scene* yang dinamis dan mempertahankan identitas objek target dengan tepat tanpa memperhatikan oklusi dan beberapa *perturbation* visual yang lain. Penjejakan objek berbasis visual mempunyai problem yang menarik dan menjadi tantangan. Problem penjejakan objek target yang *robust* dalam aplikasi penjejakan pemain *soccer* diangkat oleh peneliti [1-2]. Berbagai pendekatan telah dilakukan untuk mengeksplorasi berbagai konsep dasar penjejakan objek di antaranya yaitu pendekatan tahap-tahap umum dalam penjejakan objek. Makalah ini menyajikan hasil kajian singkat terhadap berbagai metode terkait pengendalian dan penjejakan objek yang telah dilakukan para peneliti sebelumnya melalui berbagai artikel ilmiah yang relevan dengan topik pengendalian dan penjejakan objek. Beberapa pengelompokan terhadap penelitian di bidang ini dilakukan dengan tujuan untuk memberikan gambaran dan panduan dalam melakukan penelitian-penelitian selanjutnya.

Kajian singkat terhadap penelitian di bidang pengendalian dan penjejakan objek ini disajikan dengan sistematika sebagai berikut ini. Bab 2 menyajikan tentang hal-hal mendasar terkait deteksi dan pengenalan objek. Bab 3 berisi tentang estimasi pergerakan objek. Bab 4 menjelaskan aspek pengendalian dan penjejakan objek. Terakhir, hal-hal penting yang disampaikan di makalah ini disusun sebagai kesimpulan pada Bab 5.

2. Deteksi Dan Pengenalan Pola

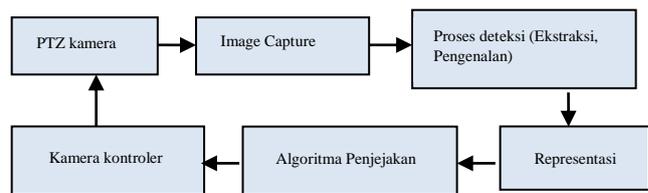
2.1 Deteksi Objek

Pendeteksian objek merupakan *task* yang penting dan menantang dalam teknologi visi komputer. Visi komputer adalah salah satu bidang aplikasi dari *artificial intelligence* yang memiliki fokus pada informasi-informasi berupa citra foto, citra bergerak (video), ataupun informasi yang didapatkan dari data multidimensional *scanner* medis. Bidang aplikasi kecerdasan buatan dalam proses pengolahan citra berfungsi seperti memperbaiki kualitas citra, pemanipulasian citra, pemampatan citra, ataupun analisis sebuah citra.

Deteksi objek adalah prosedur untuk menemukan entitas non-stasioner dalam urutan gambar. Pada *task* pendeteksian objek bertujuan untuk menghasilkan keputusan ciri yang dianggap mewakili atau merepresentasikan aspek objek atau tidak dan sebagai tahap awal untuk melacak objek yang bergerak sebelum dilakukan proses representasi dan pengenalan objek serta

sebagai langkah selanjutnya untuk menjejak. Proses mendeteksi objek ini melibatkan berbagai metode pendeteksian objek sehingga fitur atau ciri yang dihasilkan optimal. Terlihat **Gambar 1** menunjukkan blok diagram secara umum untuk deteksi objek.

Penjejakan sebagai metode identifikasi dari posisi objek yang bergerak. Mengidentifikasi posisi juga merupakan *task* yang jauh lebih menantang disamping mendeteksi benda bergerak. Penjejakan objek banyak diterapkan dalam berbagai aplikasi seperti dalam visi robot, pemantauan lalu lintas, video *surveillance*, dan simulasi.



Gambar 1: Blok diagram deteksi objek

Beberapa metode pendeteksian objek telah digunakan di dalam penelitian-penelitian tersebut sebagian besarnya adalah metode Mean Shift [3], Continously Adaptive Mean Shift (Camshift) [4-5], Adaptive Boosting Viola[6-9], Gaussian Mixture Model [10-11], PCA (Principal Component Analysis) [12], dan Epipolar Geometry [13]. Sebuah tinjauan yang dilakukan oleh Yilmaz dkk dengan memberikan gambaran klasifikasi dari deteksi dan penjejakan objek [14]. Beberapa *survey* tentang deteksi objek pada area dengan tantangan dan lingkungan yang berbeda, diantaranya ketika sistem deteksi objek dilakukan di area yang melibatkan parameter kondisi pencahayaan, parameter tubuh dan atribut pada objek [15-16]; kondisi cuaca seperti hujan, kabut dan salju; jarak pendeteksian, dan kecepatan mendeteksi. Sistem pelacakan juga ditawarkan dengan menggunakan sensor termal 3D pada ruang terbuka (non-laboratorium), termasuk harus memperhatikan variabel kondisi pencahayaan, hujan, asap, jarak pelacakan, dan kecepatan pelacakan. Pengaruh kondisi cuaca ini dapat mengacaukan sistem lain, namun objek dapat dilacak dalam 3D dengan kesuksesan tinggi. Sistem sensor termal 3D yaitu PROWL (Perception for Robotic Operation over Widespread Lighting), menggunakan pemrosesan gambar stereo termal dan pemrosesan *sensor on-board* untuk pelacakan objek dan pengenalan gerakan.

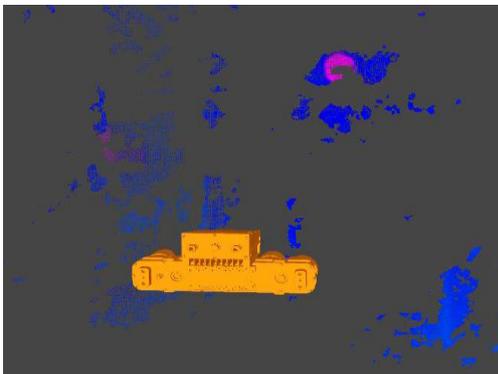
Sistem sensor PROWL menggunakan algoritma *point matching* berbasis-ICP dan berhasil mendapatkan 100% pelacakan objek pada 20 *frame* per detik hingga 13 meter dan pengenalan gerakan menghasilkan zero dalam false-positive/false-negative dengan jarak 7 meter pada

berbagai skenario uji. Skenario yang diuji dipilih pada lingkungan luar yang cerah, lingkungan luar di malam hari, lingkungan *indoor* yang gelap, dan lingkungan *indoor* yang dipenuhi kabut asap.

Sensor ini memberikan solusi kendala parameter-parameter tersebut. Sistem sensor yang digunakan adalah *thermal stereo image processing* dan *sensor processing on board*, sehingga mampu mendeteksi objek dan mengenali *gesture*. Pada **Gambar 2** menjelaskan bagaimana objek kepala manusia dapat dibedakan dari lingkungannya, saat menggunakan data termal dibandingkan dengan menggunakan *visible spectrum data*. Dalam lingkungan yang lebih hangat, data sensor berisi objek non-manusia dengan suhu yang sama dengan objek manusia, tetapi data tersebut dengan cepat dihapus karena dianggap tidak menarik selama pelacakan 3D berdasarkan bentuk atau posisi 3D objek terkait dengan frame sebelumnya. Objek wajah manusia yang sedang mengintip dari balik penghalang diperlihatkan pada **Gambar 2a**. Dengan RGB image objek wajah manusia sulit mendeteksi jarak yang diberikan dari sensor (warna *orange*) dan sulit mendeteksi gangguan lain pada area tersebut. Sementara pada **Gambar 2b**, objek wajah manusia dengan mudah dibedakan berdasarkan suhu dalam *thermal point cloud (magenta pixel)*.



Gambar 2a: Hasil pendeteksian objek dengan RGB *point cloud*



Gambar 2b: Hasil pendeteksian objek dengan *thermal point cloud*



Gambar 3a: Hasil pendeteksian dengan RGB sebelum objek masuk dalam ruangan berkabut/berasap



Gambar 3b: Hasil pendeteksian dengan RGB objek masuk dalam ruangan berkabut/berasap



Gambar 3c: Hasil pendeteksian objek dengan thermal point cloud ketika objek masuk dalam ruangan berkabut/berasap.

Gambar 3 menjelaskan perbedaan objek manusia antara dua *point cloud* dalam kondisi *whiteout* (dalam ruangan yang hangat dan berasap). Dengan menggunakan sensor RGB dan thermal dari PROWL. Lingkaran hijau pada **Gambar 3a** menunjukkan pendeteksian objek kepala manusia secara 3D sebelum objek masuk dalam ruangan berasap. **Gambar 3b** menjelaskan pendeteksian dengan sensor *visible-spectrum cloud*

terlihat bingung oleh asap yang mengaburkan objek. Sementara dengan sensor kamera *thermal point cloud* objek terdeteksi dan terlihat, serta tidak terpengaruh oleh asap, terlihat pada **Gambar 3c**.

Terdapat celah evaluasi yang dapat diangkat dari metode ini, bahwa batas sistem ini yaitu pada beberapa lingkungan, seperti ruangan hangat yang mendekati suhu tubuh manusia. Hal ini yang membatasi manfaat yang terlihat dari penggunaan sensor thermal 3D, karena lebih sulit untuk melakukan segmentasi objek pada lingkungan tersebut, artinya lingkungan tidak dapat sepenuhnya disaring oleh sensor thermal. Sebagai alternatif dengan menggunakan pelacak berbasis sensor stereo RGB kemungkinan bisa dicoba dengan menggunakan *Microsoft kinect* atau *spinning LIDAR*. Namun sensor thermal 3D dapat terjadi *false positive* dengan metode pelacakan untuk objek yang mendekati atau mengaburkan target.

Metode *viola-jones* dan *template machine* digunakan untuk pendeteksian wajah di seluruh *frame window* dan melakukan implementasi dengan perangkat webcam kamera, servo sebagai actuator, dan *arduino uno board*. Proses deteksi, sekali objek wajah ditemukan maka posisinya akan disimpan dan akan dihitung *region of interest (ROI)* di sekitar objek wajah dan digunakan untuk deteksi wajah di sekitar ROI tersebut, sehingga dari proses ini dibutuhkan waktu yang lebih lama dari proses yang pertama karena deteksi wajah hanya dilakukan di sekitar ROI wajah. Sementara metode *template matching* berperan sebagai *back-up* jika terjadi kegagalan deteksi dengan metode *viola jones*.

Problem muncul saat deteksi objek yaitu *foreground* tidak ada atau *background* selalu berubah-ubah dari setiap *scene*-nya karena pengaruh intensitas cahaya (intensitas piksel) dan mengakibatkan nilai pikselnya di luar nilai ambang [15]. Penyempurnaan dilakukan dengan menggunakan metode pembuat *background* yang dapat beradaptasi, yaitu metode PCA-GMM [10] bertujuan menghasilkan bentuk fitur yang presisi serta jelas dalam *scene* dan bisa memberikan klasifikasi yang baik terhadap objek. Segmentasi menggunakan GMM terlihat pada **Gambar 4a**. Sementara untuk mengatasi masalah pelacakan menggunakan segmentasi PCA-GMM dan mencapai peningkatan yang luar biasa, ditunjukkan pada **Gambar 4b**.

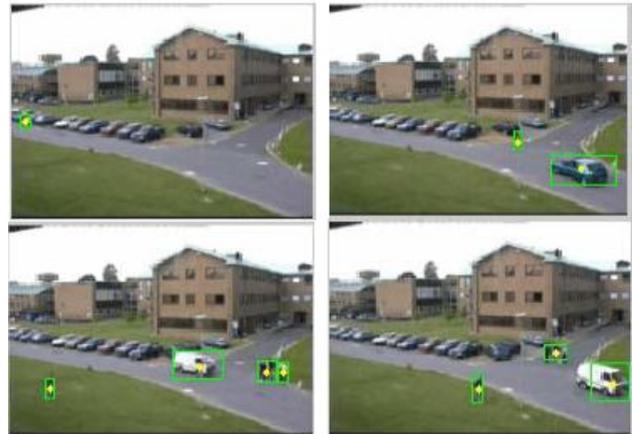
Kemudian penjejakan pada objek bergerak berbasis KF diimplementasikan dengan menggabungkan metode *principal component analysis-gaussian mixture method-kalman filter (PCA-GMM-KF)*. Hasil penerapan metode PCA-GMM-KF menunjukkan penjejakan terhadap objek yang lebih akurat, ditunjukkan pada **Gambar 5a** dan **Gambar 5b**.



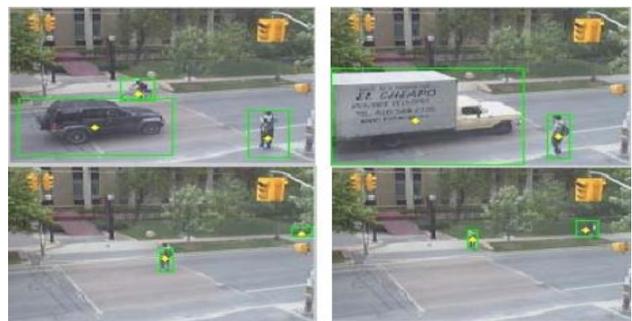
Gambar 4a: Segmentasi objek bergerak dengan algoritma GMM



Gambar 4b: Segmentasi objek bergerak dengan algoritma PCA-GMM



Gambar 5a: Segmentasi dari objek yang bergerak pada area 1 dengan algoritma penggabungan PCA-GMM-KF



Gambar 5b: Segmentasi dari objek yang bergerak pada area 2 dengan algoritma penggabungan PCA-GMM-KF

2.2 Pengenalan Objek

Pengenalan wajah yaitu membandingkan citra wajah masukan dengan suatu *database* wajah dan menemukan wajah yang paling cocok dengan citra masukan tersebut. Pengenalan objek berkaitan dengan biometrik yaitu pengenalan identitas melalui karakter visi. Pengenalan objek secara umum terdiri atas sub-sub proses yang sama dengan pengenalan pola pada data umum lainnya, yang bisa dimodelkan yaitu sumber cahaya, objek, dan sensor. Ketiga komponen ini sangat penting dalam pembentukan sebuah persepsi visual.

Pemahaman data *image* juga mempunyai masalah tersendiri untuk pengenalan pola. Pergerakan objek *image* dapat dianggap sebagai transisi pola nilai piksel dari waktu ke waktu yang dapat ditunjukkan oleh fungsi transformasi. Dengan pendekatan metode pengenalan objek juga dilakukan dengan menggunakan algoritma *individual skeletal features*, dengan metode ekstrak berbasis *principal component analysis (PCA)*, dan menggunakan *kinect sensor on-board* pada *mobile robot* [12]. Penelitian ini fokus pada (i) penggunaan *kinect* sehingga menekan harga dari sistem monitor *care robot (robot mobile)* dengan ukuran yang *compact*, (ii) dilakukan terhadap objek berdasarkan perbedaan skeleton dengan *rectangular color marker*. *Care robot* ini untuk merealisasikan upaya mengatasi problem yang diangkat (yaitu menjaga *safety life*), dengan kemampuan antara lain pengenalan (*recognizing*), pelacakan (*tracking*), dan *monitoring*. Pelacakan *skeletal* dengan *kinect* menghasilkan posisi 20 titik *skeletal* yang berbeda, kemudian fitur objek orang diekstrak dan dihitung dengan *euclidean norm*. Namun metode ini masih ada keterbatasan *noise* dari pengukuran *depth* pada *recognition* dan *tracking* yang sesekali tidak stabil. Andriana dkk mencoba memberikan solusi dengan mendekati struktur pola sinyal data dalam deret waktu [15]. Dengan mencoba menggunakan metode hibrida yang menggabungkan metode statistik untuk membedakan objek dan metode struktural untuk menggambarkan fitur morfologis (berbasis bentuk). Biasanya, bentuk benda dalam ekstraksi fitur morfologi digambarkan dengan tepi, lekukan, dan daerah. Fitur teknik ekstraksi meliputi kode rantai, regresi linear *piece-wise*, dan *curve fitting*. Regresi linier berganda pada integral *choquet* juga dicoba untuk memperbaiki pengenalan wajah, untuk mempelajari tepi *sobel*, dan gradien morfologis. Namun masih memerlukan lebih banyak sampel pengujian dan *training*. Dan operator morfologi tingkat abu-abu seperti pelebaran dan erosi dicoba digunakan, juga tetap tidak menunjukkan rincian lebih lanjut tentang karakteristik bentuk dan fungsi transformasinya. Perbandingan metode pengujian yang dilakukan ditunjukkan pada **Tabel 1**.

Tabel 1: Perbandingan beberapa metode yang dilakukan dari *area of interest*

Metode / Texture Description	Area of Interest	Citra	Hasil
MRL	Sekitar mata	Gray level	Akurasi posisi mata & alis dari 2 orang yang berbeda
	Bahu	Contour	Koefisien determinasi lebih kecil, standar deviasi tetap
LBP	Sekitar mata	Gray level	Nilai X^2 terkecil (identifikasi yg benar)
SURF	Postur tubuh	Siluet & Contour	False positive tinggi menjadi rendah
HOG-SVM	Postur tubuh	Contour	True positive (correctness of classification) & Zero false positive (zero wrong classification)

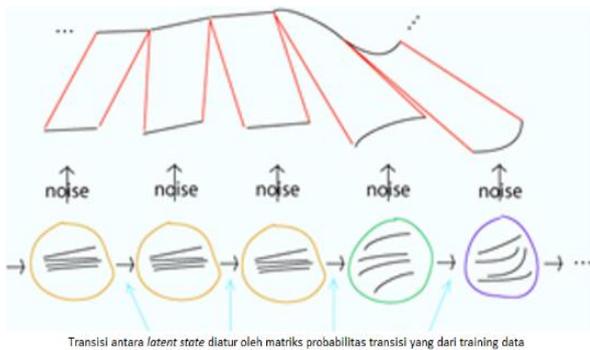
Pengujian dilakukan kembali dengan mengusulkan pengkombinasian dari fitur wajah dan *posture* untuk penjejakan karakteristik gerak objek visual. Dengan pendekatan *curve fitting* pada fitur wajah dan *posture* tubuh dengan menangkap pola fitur dengan tepat. Pada *posture* tubuh dilakukan *pre-processing* dengan kamera *kinect depth* dan membandingkan metode pengenalan objek visual. Namun performansi menggunakan seluruh wajah atau seluruh area tubuh bisa berkurang karena pergerakan pada pose yang berbeda. Untuk mengenali dan membedakan karakteristik wajah dan tubuh secara lebih baik, Andriana dkk mengembangkan lagi dengan menggunakan LBP untuk area yang diidentifikasi lebih kecil seperti sekitar mata, hidung, mulut, tekstur rambut, jilbab, atau topi. Metode yang diusulkan digunakan untuk klasifikasi objek visual. Probabilitas dari identifikasi diri ditingkatkan dengan menggunakan *pose* dan karakteristik yang berbeda dari detail fitur yang lebih kecil pada area *pose* tubuh dan wajah.

3. Estimasi Pergerakan Objek

Estimasi pergerakan objek merupakan proses untuk menentukan pergerakan objek pada video secara berurutan. Pergerakan obyek dikenal dengan istilah *motion vector*. *Motion vector* menunjukkan pergeseran sebuah titik diantara *frame* sekarang dengan *frame*

referensi. Dari *motion vector* yang didapat, akan terlihat pergerakan titik-titik yang terjadi antara *frame* yang diamati. Teknik estimasi pergerakan objek sudah dilakukan oleh beberapa peneliti seperti teknik *pixel based motion estimation* dan *block matching algorithm* [16]. Vasquez dkk melakukan *clustering* hierarkis dengan metrik *euclidean* untuk mempartisi lintasan pelatihan ke dalam *cluster*, mempelajari cara dan varian dari lintasan dalam *cluster*, dan memprediksi dengan menggunakan lintasan rata-rata dari *cluster* yang paling mungkin dari masukan segmen lintasan. Piciarelli dan Foresti membangun *tree of cluster* dengan meminimalkan jarak *euclidean* yang dinormalisasi, dan menggunakan *filter kalman* untuk prediksi lintasan [18].

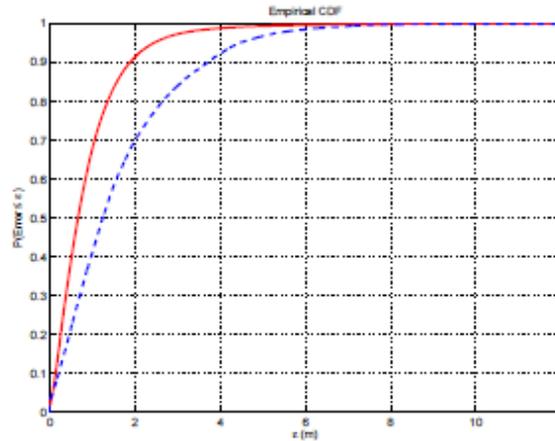
Choi dan Hebert melakukan pendekatan pergerakan *short term behavior* dengan pemodelan trajektori sebagai rangkaian segmen pendek, diasumsikan sebagai *noise* dari segmen *latent*, dan menggunakan model *markov* dari segmen trajektori *piecewise* [19]. Pendekatan untuk memprediksi *future motion* objek berbasis pada pergerakan objek sebelumnya. Diperlihatkan pada **Gambar 6** bahwa jumlah *latent state* dapat mempengaruhi kinerja model prediktif atau estimasi. Terlalu banyak *latent state* akan menghasilkan beberapa *latent state* dengan titik data yang sangat sedikit sehingga menyebabkan lintasan *latent state* ini akan memiliki varian yang tinggi. Terlalu sedikit *latent state* akan menyebabkan model gagal untuk menangkap struktur segmen lintasan.



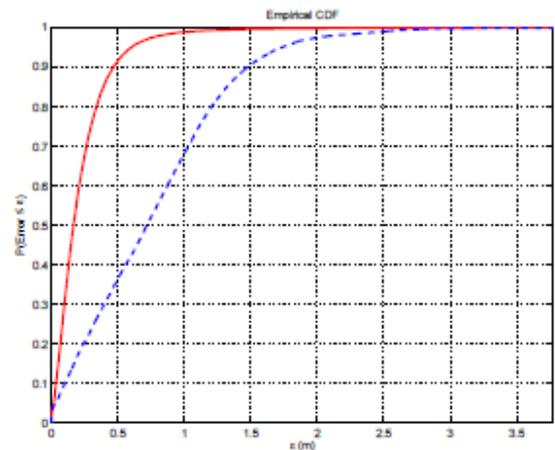
Gambar 6: Model trajektori objek: hubungan antara *observed trajectory*, *observed trajectory segment*, dan *segmen latent*

State dengan kondisi yang lebih *latent* akan memperlambat model. Pendekatan algoritma *k-means* untuk memilih jumlah *cluster* dengan kriteria BIC dapat memecahkan masalah ketika memilih jumlah *latent state*. Namun, tetap tidak ada jaminan bahwa jumlah *cluster* yang ditemukan dengan kriteria BIC akan menghasilkan kinerja prediktif terbaik. Pada **Gambar 7a** dan **Gambar 7b** menunjukkan evaluasi dataset PAT bus untuk distribusi kumulatif fungsi plot dari *error* prediksi

empiris dibandingkan dengan *ground truth trajectory*. Dari setiap plot, probabilitas sumbu y dari model prediksi membuat error kurang dari *threshold* sumbu x. Garis putus-putus biru mewakili model *markov* orde pertama; Garis *solid* merah mewakili model *markov* orde dua. Dominasi garis *solid* merah di atas garis putus-putus biru, menunjukkan bahwa model *markov* orde dua lebih mungkin daripada model *markov* orde pertama untuk membuat prediksi yang mendekati lokasi target.



Gambar 7a: Prediksi pada 1 detik ke depan



Gambar 7b: Prediksi pada 2 detik ke depan

Penelitian-penelitian yang berfokus pada problem penjejakn mengalami perkembangan dan menyoroti dari berbagai bidang seperti olahraga *soccer*, *hockey*, sistem pengawasan, dan aplikasi lainnya, serta algoritma *artificial intelligence*, teknik sinyal *filtering*, kemampuan kendali, dan sistem visual. Bishop dan Spong mencoba mengangkat problem penjejakn pada sistem permainan robot *hockey* dengan mengadopsi rule permainan dan peralatan robot [20]. Aplikasi ini dilakukan pada skala laboratorium dengan menggunakan kontroler CataLyst-5 manipulator dari CRS *robotics corporation* dan kontroler C500C. Pada kontroler C500C mengimplementasikan strategi kendali visual, yang bertujuan memberikan

estimasi yang tepat dan akurat dari trajektori, lokasi, dan prediksi posisi. Pendekatan dilakukan dengan memodelkan keragu-raguan (*uncertainties*) dan kesalahan pengukuran (*measurement errors*) sebagai *white noise*. Dengan mengambil prediksi posisi menggunakan pendekatan *kalman filter*. Park dkk mengusulkan prediksi berbasis *neural network*, tetapi pada *neural network* menetapkan estimasi yang akurat untuk *intercept state* akhir jika data *training set* dari trajektori yang digunakan, semua parameter *state* dimasukkan sebagai masukan dan dengan *network* mempunyai beberapa *node* internal.

Tabel 2: Evaluasi pada PAT bus dataset: 50%, 90%, 95% dari *prediksi error* pada model Markov *first-order* (MM(1)) dan model Markov *second-order* (MM(2)). Test set secara random dari model Markov (MM(1)): 4830 sample trajektori; MM(2): 3209 sample trajektori.

Seco nds into furur e	Percent of Prediction Error					
	50%		90%		95%	
	MM (1)	MM (2)	MM (1)	MM (2)	MM (1)	MM (2)
1	0.72m	0.17m	1.47m	0.47m	1.74m	0.61m
2	1.25m	0.67m	3.73m	1.89m	4.46m	2.45m
3	2.30m	1.56m	7.97m	4.47m	9.61m	5.78m
4	3.93m	2.87m	13.6m	8.33m	16.6m	10.8m
5	6.13m	4.64m	20.7m	13.5m	25.1m	17.4m
6	8.90m	6.86m	28.8m	19.9m	35.3m	25.7m
7	12.2m	9.57m	38.1m	27.6m	46.5m	35.6m
8	15.9m	12.7m	48.6m	36.3m	59.0m	47.1m
9	20.2m	16.3m	59.7m	46.1m	73.2m	60.3m
10	24.6m	20.2m	71.6m	56.9m	88.6m	78.8m

Ketika objek berpindah dan saling berinteraksi satu dengan yang lain memperlihatkan sebuah trajektori, kemudian memberikan relasi satu dengan yang lainnya. Relasinya berupa tipe dasar informasi untuk menentukan trajektori pada penjejakan selanjutnya. Model berbasis *relationship* antara trajektori objek. Ketika mengekstrak informasi diperlukan pengukuran kinerja dari interaksi antara pemain dalam satu tim dan tim lawan, (i) perlu memperluas model dengan mengekstrak informasi trajektori, sehingga ada peluang untuk memperluas penelitian di bidang problem trajektori untuk diterapkan pada permainan soccer dan permainan olahraga lain yang masih menggunakan bola, seperti *handball*, *hockey*, dan *basket ball*. (ii) model berbasis *relationship* antara *event* penting seperti *goal* dan *performance measure* dengan data *real*.

4. Pengendalian Dan Penjejakan Objek

Masalah terkait pengendalian dan penjejakan objek dilakukan dengan pendekatan implementasi kamera statik

dan *pan-tilt-zoom* (PTZ) berbasis sistem penjejakan objek secara *real-time*. Proses penjejakan objek yang terdeteksi oleh kamera, bertujuan untuk menentukan koordinat pusat dari wajah dan kemudian koordinat tersebut digunakan untuk menggerakkan sistem *pan-tilt-zoom*, sehingga objek akan selalu berada di tengah *frame* atau *field of view* (FOV) [7-8]. Namun sistem penjejakannya masih tidak bekerja ketika ada lebih dari satu objek yang terdeteksi. Lee dan Xiong mengimplementasikan pergerakan kamera yang dikendalikan oleh *command* dalam bentuk *single* 3 sampai 16 *bytes* dan sinyal ACK yang dihasilkan akan mentransmit kembali ke *controller* [7] dan tampilan implementasinya ditunjukkan pada **Gambar 8**. Yosafat dkk mengembangkan dan mengimplementasikan sistem pengendalian dengan membandingkan antara kendali *lead-lag* dan kendali PID [8].



Gambar 8: Tampilan depan implementasi sistem kamera PTZ

Namun hasil uji implementasinya masih terdapat beberapa keterbatasan yaitu ketika sistem mendeteksi lebih dari satu objek wajah, maka sistem tidak berjalan dengan baik karena tidak dapat memutuskan wajah yang harus diikuti. Kedua sudut dan kemiringan wajah terhadap kamera, ketika wajah tidak menghadap persis dan tidak tegak lurus ke kamera (ada batasan sudut kemiringan objek wajah), sistem tidak mendeteksi keberadaan wajah karena perbedaan tingkat intensitas tidak dapat dihitung dengan satu sisi objek wajah yang terlihat sehingga tidak mendapatkan variasi pada tingkat intensitas. Semakin besar sudut kemiringan, semakin sulit juga sistem untuk mendeteksi, karena hanya memetakan objek wajah standar yaitu bagian depan wajah (*frontal face*). Ketiga adalah faktor jarak objek terhadap kamera, sistem deteksi wajah didesain berada pada jarak kurang dari 3 meter. Artinya wajah yang memiliki jarak paling dekat ke kamera dan dengan ukuran yang terbesar yang terdeteksi oleh sistem. Jika semakin jauh wajah dari kamera, maka semakin kecil ukuran objek tersebut dalam *frame window*. Sehingga ukuran yang kecil dari objek wajah menyulitkan fitur *haar* untuk menghitung perbedaan intensitas antara daerah yang gelap dan terang pada objek wajah. Namun semakin besar ukuran *frame*

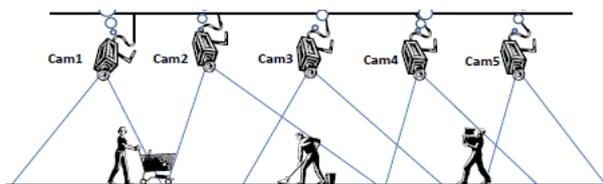
yang menampilkan objek wajah, menyebabkan semakin banyak waktu yang dibutuhkan untuk mendeteksi keberadaan wajah. **Tabel 3** menunjukkan pengaruh ukuran *frame* terhadap waktu deteksi.

Tabel 3: Pengaruh ukuran *frame* terhadap waktu deteksi

	Ukuran frame				
	240x 160	320x 240	640x 480	800x 600	1024x 800
Waktu (detik)	0.011	0.013	0.016	0.047	0.082

Keterbatasan keempat adalah beberapa oklusi diberikan pada bagian mata objek wajah, dengan pemakaian kacamata dan masker, sistem masih dapat mendeteksi karena perbedaan intensitas pada bagian mata dan hidung masih menghasilkan variasi yang cukup untuk mendeteksi wajah. Tetapi ketika mata yang ditutupi oleh tangan, sistem tidak mendeteksi dengan baik, karena algoritma *viola-jones* mementingkan bagian mata dan hidung dalam proses deteksi dan pada bagian mata perbedaan intensitas daerah gelap dan terang cukup tinggi.

Sementara Natarajan, mengusulkan pendekatan teori keputusan untuk koordinasi pengendalian beberapa kamera aktif (*pan-tilt-zoom*) dan untuk mencapai pengawasan dengan kualitas tinggi [22]. Pendekatan teori keputusan dimodelkan sebagai fungsi objektif *robust mathematical* yaitu pemodelan yang saling berinteraksi antara jaringan kamera aktif dan lingkungan pengawasan. Pendekatan teori keputusan dilakukan sebagai perencanaan pengendalian secara optimal dari *task* pengawasan untuk kamera aktif pada gerak objek, lokasi objek, dan lain-lain yang berubah-ubah; jumlah objek yang lebih banyak daripada jumlah kamera yang tersedia; *framework* koordinasi kamera dapat ditingkatkan untuk menambah jumlah objek target. Sementara untuk pengawasan secara *real-time* dilakukan dengan pengaturan mekanikal ditunjukkan pada **Gambar 9** yang mengkoordinasi pengendalian pada beberapa kamera.



Gambar 9: Visualisasi koordinasi pengendalian dari beberapa kamera

Beberapa permasalahan terkait dengan aplikasi kajian pengendalian dan penjejakan objek dapat diselesaikan dengan beberapa pendekatan, dipaparkan capaian yang telah ditempuh oleh penelitian sebelumnya

dalam bidang pendidikan, teknologi penjejakan objek dimanfaatkan pada *platform* pendidikan jarak jauh (*online*) secara *real-time*. Bidang kesehatan dan kesejahteraan, Cao dan Hashimoto merancang *care robot* dengan ukuran yang sesuai dan biaya rendah, yang difokuskan untuk menjaga *safety life* dengan kemampuan pengenalan (*recognizing*), penjejakan (*tracking*), dan *monitoring* [12]. Metoda pelacakan disempurnakan secara *real time* dan pengembangan teknik segmentasi untuk proses pengenalan dan mencegah estimasi yang salah, diaplikasikan pada bidang pengawasan di supermarket, dengan fokus pelacakan untuk jumlah objek manusia yang banyak (*multi person tracking*)[23]. Sistem pelacakan dikembangkan dengan pendekatan sistem *distributed realtime* untuk melacak lokasi objek dengan menggunakan *multiple calibrated camera* secara *real time* dan Krumm dkk menggunakan dua *set* kamera *stereo* untuk penjejakan objek pada *indoor environment* (seperti ruang tamu, ruang kelas, dan ruang konferensi) [25]. Pada teknologi aplikasi penjejakan digunakan untuk semua orang yang ingin berkinerja lebih baik antara lain untuk *divers jumping* untuk mengatur jejak kaki dan menterjemahkan lompatannya yang tinggi dan target waktu untuk melakukan gerakan berputar (*twist*); pelari dengan posisi *tracking*; dan pemanah dengan posisi atau kondisi stabil untuk mencapai target yang baik [26].

5. Kesimpulan

Dari uraian di atas dapat disimpulkan bahwa objek yang menjadi target untuk deteksi dan pengenalan adalah manusia dalam kondisi bergerak dengan pengamatan melalui pergerakan pada *image* sekuensialnya, dengan karakteristik wajah dan *pose* tubuh yang dijadikan informasi. Sistem penjejakan akan mengendalikan gerakan horisontal dan vertikal dari kamera agar posisi manusia yang dikenali selalu berada di tengah *frame* kamera.

Pengendalian dan penjejakan pada objek yang bergerak bertujuan untuk menghasilkan informasi yang lebih akurat. Sehingga beberapa capaian yang telah ditempuh oleh penelitian-penelitian sebelumnya mencakup dari berbagai bidang aplikasi.

Penjejakan dengan melibatkan banyak *sensors* pada objek, menjadi peluang untuk penelitian terkait dengan (i) penjejakan yang berbeda dengan objek target yang sama, (ii) metoda mengkombinasikan *state* estimasi untuk sistem penjejakan, (iii) aspek pengembangan keakuratan dari sensor atau kinerja sensor yang sama tetapi dengan harga yang lebih murah, dan (iv) aspek sistem *embedded* yang akan diimplementasikan dengan pengendalian dan penjejakan objek secara *realtime operating system* pada beberapa aplikasi.

Daftar Pustaka

1. C-H. Kan^o, I-R Hwan^o, dan K-J Li, *Trajectory Analysis for Soccer Players*. Proceeding ICDMW '06 Proceedings of the Sixth IEEE International Conference on Data Mining–Workshops, Washington, USA, December 18 - 22, 2006, 377-381, ISBN: 0-7695-2702-7.
2. J. Kang, I. Cohen dan G. Medioni. *Soccer Player Tracking across Uncalibrated Camera Streams*, IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (PETS.03) In conjunction with ICCV, Nice, France. October, 2003.
3. X. Lu, J. Zhang, L. Song, R. Lei, H. Lu, dan N. Ling. *Person Tracking with Partial Occlusion Handling*, 2015 IEEE Workshop on Signal Processing Systems (SiPS). Hangzhou, China. October 14-16, 2015, 1-6, ISBN: 978-1-4673-9604-2
4. A. Mian, *Realtime Face Detection and Tracking Using a Single Pan, Tilt, Zoom Camera*, International Conference on Advances in Computing, Communications and Informatics (ICACCI). Christchurch, New Zealand. November 26-28, 2008, 1-6, ISBN: 978-1-4244-2582-2.
5. Z. Tang, Y-S. Lin, K-H Lee, J-N Hwang, J-H Chuang, dan Z. Fang. *Camera Self-Calibration from Tracking of Moving Persons*. 23rd International Conference on Pattern Recognition (ICPR) Cancun Center. Cancun. Mexico. December 4-8, 2016, 265-270, ISBN: 978-1-5090-4847-2.
6. A. Kapusta dan P. Beeson. *Person Tracking and Gesture Recognition in Challenging Visibility Conditions Using 3D Thermal Sensing*, 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Columbia University, NY. USA. August 26-31, 2016, 1112-1119, ISBN: 978-1-5090-3929-6.
7. S. Lee dan Z. Xiong. *A Real-time Face Tracking System Based On A Single Ptz Camera*, Signal and Information Processing (ChinaSIP), IEEE China Summit and International Conference. Chengdu. China. July 12-15, 2015, 568-572, ISBN: 978-1-4799-1948-2.
8. R. Yosafat, C. Machbub dan E. M. I. Hidayat, *Design and Implementation of Pan-Tilt Control for Face Tracking*, IEEE International Conference on System Engineering and Technology. Shah Alam. Malaysia. October 2-3, 2017, 217-222, ISBN: 978-1-5386-0383-3.
9. X. Li, K. Wang, W. Wang dan Y. Li. *A Multiple Object Tracking Method using Kalman Filter*, IEEE International Conference on Information and Automation (ICIA). Harbin. China. June 20-23, 2010, 1862-1866, ISBN: 978-1-4244-5704-5.
10. E. Noureldaim, M. Jedra dan N. Zahid. Multiple Tracking of Moving Objects with Kalman Filter and PCA-GMM Method, *Intelligent Information Management*, 5(2), 29590, 2013.
11. D. Focken, dan R. Stiefelhagen. *Towards Vision-based 3-D People Tracking in a Smart Room*, Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces (ICMI'02). Pittsburgh, PA. USA. October 16-16, 2002, 400-405, ISBN: 0-7695-1834-6.
12. M. Cao dan H. Hashimoto. *Specific Person Recognition and Tracking of Mobile Robot with Kinect 3D Sensor*, Industrial Electronics Society, IFCON 2013-39th Annual Conference of the IEEE. Vienna. Austria. November 10-13, 2013 8323-8328, ISBN: 978-1-4799-0224-8.
13. E. Konstantinou dan I. Brilakis. *3D Matching of Resource Vision Tracking Trajectories*. Conference: Construction Research Congress. San Juan. Puerto Rico. May 31–June 2, 2016, 1742-1752, ISBN: 9780784479827.
14. A. Yilmaz, O. Javed dan M. Shah. Object Tracking: A Survey, *ACM Computing Surveys*, 38(4), 1-45, 2006.
15. D. Andriana, A. S. Prihatmanto, E. M. I. Hidayat, I. Supriana, C. Machbub. Combination of face and posture feature for tracking of moving human visual characteristics, *International Journal on Electrical Engineering and Informatics*, 9(3), 616-631, 2017.
16. I. E. G. Richardson, *H.264 And MPEG-4 Video Compression*, England: John Wiley & Sons Ltd, 2003.
17. D. Vasquez, Th. Fraichard, and C. Laugier. Growing Hidden Markov Models: a Tool for Incremental Learning and Prediction of Motion. *International Journal of Robotics Research*, 28(11), 71-82, 2009.
18. C. Piciarelli, G. L. Foresti. On-line trajectory clustering for anomalous events detection *Pattern Recognition Letters*, 27(15), 1835-1842, 2006.
19. P. P. Choi dan M. Hebert. M. Learning and predicting moving object trajectory: a piecewise trajectory segment approach. *Robotics Institute*, 337, 1-17, 2006.
20. B. F. Rishon dan M. W. Spong. Vision-based control of an air hockey playing robot. *IEEE Control Systems*, 19(3), 23-32, 1999.
21. P. H. Shilpa dan M. R. Sunitha, A Survey on Moving Object Detection and Tracking Techniques. *International Journal Of Engineering And Computer Science*. 5(5), 16376-16382, 2016.
22. P. Natarajan, P. K. Atre, M. Kankanalli. Multi-Camera Coordination and Control in Surveillance Systems: A Survey. *ACM Trans. Multimedia Comput. Commun. Appl.*, 11(4), 1-57, 2015.
23. C. Migniot, dan F. Ababsa. *Part-based 3D multi-person tracking using depth cue in a top view*, International Conference on Computer Vision Theory and Applications (VISAPP). Lisbon. Portugal. January 5-8, 2014, 419-426, ISBN: 978-9-8975-8133-5.
24. T. Ambrish, K. Mark, W. D. James dan P. Gerasimos. Method for 3D Tracking Using Multiple Cameras, *OSU Dept. of Computer Science and Engineering Technical Report*, 2006.
25. J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale S. Shafer, *Multi-Camera Multi-Person Tracking for EasyLiving*, Proceedings. Third IEEE International Workshop on Visual Surveillance. Dublin, Ireland, July 1, 2000, 3-10, ISBN: 0-7695-0698-4.
26. S. Haake, The impact of technology on sporting performance in Olympic sports, *Journal of Sports Sciences*, 27(13), 1421-1431, 2012.
27. J. L. Park, C. B. Partridge, dan M. W. Spong, Neural Network Based State Prediction for Strategy Planning of an Air Hockey Robot, *Journal of Robotic Systems*, 18(4), 187-196, 2001.