Obstacle Avoidance Using Image-based Visual Servoing Integrated with Nonlinear Model Predictive Control

Daewon Lee∗, Hyon Lim† and H. Jin Kim‡
Seoul National University, Seoul, Republic of Korea
{dwh001∗, hyonlim†, hjinkim‡}@snu.ac.kr

Abstract—This paper proposes a vision-based obstacle avoidance strategy in a dynamic environment for a fixed-wing unmanned aerial vehicle (UAV). In order to apply a nonlinear model predictive control (NMPC) framework to image-based visual servoing (IBVS), a dynamic model from UAV control input to image features is derived. From this dynamics, a visual information-based obstacle avoidance strategy in an unknown environment is proposed. When a vision system is employed on a UAV, it is easy to lose visibility of the target in the image plane due to its maneuvering. To address this issue, a visibility constraint is considered in the NMPC framework. The advantage of the proposed method is that the constraints (e.g., visibility maintaining, actuator saturation) can be modeled and solved in a unified framework. Numerical simulations on a UAV model show satisfactory results in reference tracking and obstacle avoidance maneuvers with the constraints.

I. INTRODUCTION

Cameras have become ubiquitous in unmanned aerial vehicle (UAV) systems. Vision allows a UAV to obtain not only surveillance images but also qualitative information on the surrounding environment. Therefore, many researchers have exploited this sensor for various purposes beyond surveillance. In particular, image-based visual servoing (IBVS) has been actively studied [1], [2], [3], because the sensor gives rich information about the environment with low-cost compared to radar systems.

The objective of IBVS is to control the system to place the target image features in the desired image position. To this end, classical IBVS approaches utilize an image Jacobian matrix which relates image plane velocity to camera body velocity [1]. The image Jacobian matrix is parameterized by intrinsic parameters of the camera and the depth of the feature [4]. Due to the fact that the image Jacobian matrix is parameterized by depth and pixel coordinates only, if some targets move out of the camera field of view (FOV) during the operation, the value of current features can no longer be computed. To overcome this disadvantage, IBVS with visibility maintenance has been proposed using nonlinear model predictive control (NMPC) [5], [6]. The NMPC-based IBVS allows the controller to consider visibility constraints naturally in the nonlinear optimization framework. The advantage of this approach is that there is no need to maintain control input separately for the UAV to see a specific target. The control input is generated based on the imposed constraints and predicted states over a receding horizon in order to track the desired trajectory.

There are a few works related to IBVS in the MPC framework for fully-actuated robotic manipulators. In [5], a solution to control a 6-DOF robotic manipulator to see desired image coordinates is proposed based on NMPC. Using a cost function based on errors in the image plane, convergence and stability of the robot motion have been obtained. It is shown to be robust to calibration error and measurement noise. In [6], visual predictive control (VPC) is used to move a robotic manipulator while maintaining visibility constraints. They consider mechanical and visibility constraint in the NMPC framework.

Recent studies have reported on MPC for obstacle avoidance. [7] considers obstacle avoidance for an unmanned ground vehicle (UGV) with MPC. The reference trajectories are computed for collision avoidance, without consideration of maintaining visibility of the target. For the UAV obstacle avoidance, [8] and [9] present extended Kalman filter-based target estimation method. In these papers, the reference tracking and obstacle avoidance controls are constructed in separate structures. Therefore, it is difficult to maintain the optimality in both obstacle avoidance and visibility maintenance maneuvers. As far as we know, there has been no report to avoid obstacles while maintaining visibility constraints of the target for UAV.

In this paper, a visual information-based obstacle avoidance algorithm for a fixed-wing UAV is proposed. For the UAV, three constraints should be considered in practical system: input saturation, visibility maintenance and obstacle avoidance. In order to deal with the constraints in a unified manner, the NMPC framework is applied to the control
system. Also, dynamics from UAV control inputs to image
feature states defined in the image frame is derived to
implement the visual information-based NMPC framework
simultaneously. To our knowledge, this is the first work on
obstacle avoidance algorithm for a UAV while considering
target visibility, obstacle constraints and input constraints.
The overall system flow is depicted in Fig. 2

This paper is organized as follows. In section II, a dynamic
model of the UAV and image are introduced and integrated
to derive UAV-to-image dynamics. In section III, IBVS with
NMPC for obstacle avoidance is presented. The results and
conclusion are presented in section IV and V.

II. DYNAMIC MODELS

In order to apply the NMPC framework to the IBVS for the
UAV model, the dynamics from the control input of the UAV
to the image feature must be clarified. In this section, the
dynamics of the UAV in the 3-D Cartesian coordinate frame
and the dynamics of the image features in the 2-D image
coordinate frame are introduced, briefly. Then, the integration
of both dynamics for IBVS is performed.

A. A Fixed-Wing Dynamics

A UAV model considered in this paper is shown in (1). The
parameters are based on fixed-wing UAV namely SNUACE
in laboratory as shown in Fig. 3.

\[
\dot{x}_v = f_v(x_v, u_v) = \\
\begin{bmatrix}
\dot{V} = 0 \\
\dot{\gamma} = \left(\frac{u_\gamma - g \cos \gamma}{V}\right) / V \\
\dot{\psi} = u_\psi / (V \cos \gamma) \\
\dot{X}_v = V \cos \gamma \cos \psi \\
\dot{Y}_v = V \cos \gamma \sin \psi \\
\dot{Z}_v = -V \sin \gamma
\end{bmatrix}
\]  

(1)

where \( x_v = [V \ \gamma \ \psi \ X_v \ Y_v \ Z_v]^T \) and \( u_v = [u_\gamma \ u_\psi]^T \).

Let us substitute (2) into (3), and let \( \dot{r} \) denote the velocity
of the camera:

\[
\dot{r} = \begin{bmatrix}
T \\
\Omega
\end{bmatrix} = \\
\begin{bmatrix}
T_x, T_y, T_z, \omega_x, \omega_y, \omega_z
\end{bmatrix}^T.
\]  

(4)

Then the image feature dynamics can be presented in the
following vector form:

\[
\dot{p} = J_p \dot{r}.
\]  

(5)

\( V, \gamma \) and \( \psi \) denote velocity, flight path angle and heading
angle of the UAV, respectively. \( X_v, Y_v \) and \( Z_v \) represent
position of the UAV defined in the inertial coordinate frame.
In many fixed-wing UAV operations, the velocity of the UAV
is set to be constant [10]. Therefore, the rate of the velocity
of the UAV is considered zero as shown in the first line of
(1). Control inputs \( u_\gamma \) and \( u_\psi \) denote acceleration for pitch
and yaw direction, respectively.

B. Image Dynamics

Visual servoing strategies require understanding of camera
geometry which maps the world coordinate into the image
plane. In this study, a pinhole camera model is considered
to describe the camera geometry as shown in Fig. 5.

Let \( [X_I, Y_I, Z_I]^T \) be the axes of the inertial coordinate
frame \( I \) representing the three-dimensional space, and
\( [X_c, Y_c, Z_c]^T \) be the axes of the camera frame \( C \), i.e., the
coordinate frame attached to the camera center \( O_c \). Let
\( [U_s, V_s]^T \) be the coordinate axes of the image frame \( S \),
and \( O_s \) denote the principal point where the \( z \) axis of the
camera coordinate intersects the image plane. \( f \), the distance
between \( O_c \) and \( O_s \), is the focal length of the camera.
The mapping from a point \( P = [X_c, Y_c, Z_c]^T \) in the
three-dimensional space to a point \( p = [U_s, V_s]^T \) can be written
as (2).

\[
p = \begin{bmatrix}
U_s \\
V_s
\end{bmatrix} = \begin{bmatrix}
f \\
Z_c
\end{bmatrix} \begin{bmatrix}
X_c \\
Y_c
\end{bmatrix}
\]  

(2)

Suppose that the camera is moving so that the point \( P \)
moves with translational velocity \( T = [T_x, T_y, T_z]^T \) and
rotational velocity \( \Omega = [\omega_x, \omega_y, \omega_z]^T \), with respect to
the camera frame \( C \). The dynamics of the point \( P \) with respect
to \( C \) can be expressed as (3).

\[
\dot{P} = -\Omega \times P - T
\]  

(3)
After solving \( x_r(t) \), the control input \( u(t) \) should be computed that allows state \( x(t) \) to follow \( x_r(t) \) as the following formulation:

\[
(\text{Design}) \quad u(t) \quad (12)
\]

such that  
\[
\| x(t) - x_r(t) \| \leq \varepsilon \quad (13)
\]

and  
\[
g(x(t), u(t)) \leq 0 \quad (14)
\]

where \( u(t) \) is the input, \( \varepsilon \) is a small positive real number and \( g(\cdot) \leq 0 \) are the set of constraints.

Using the visual information from a camera and a range sensor, it is difficult to generate a followable trajectory for a UAV in an unknown environment. With the single camera and range sensor information, determining the position of the multiple dynamic obstacles requires heavy computational load, and sometimes causes inaccuracy. The image-based visual servoing is applied to implement the obstacle avoidance maneuver because it does not require 3-D reconstruction. However, during the maneuver, the image feature must be kept in the image plane, otherwise there is no information with which to compute the reference trajectory. To maintain the visibility, the UAV-to-image dynamics and line-of-sight constraints should be considered simultaneously. To fulfill this requirements, the following constraints are considered in this paper.

\[
(\text{Design}) \quad u(t) \quad (15)
\]

such that  
\[
\| l_o \| - l_{sat} \leq 0 \quad (16)
\]

\[
\| u(t) \| - u_{sat} \leq 0 \quad (17)
\]

\[
1/\| P_o - P_c \| \leq \varepsilon \quad (18)
\]

and  
\[
g(x(t), u(t)) \leq 0 \quad (19)
\]

where \( l_o \) is line-of-sight (LOS) angle of an object, \( \varepsilon \) is a small positive real number, \( l_{sat} \) and \( u_{sat} \) are saturation values of the LOS and the control input, respectively. \( P_c \) and \( P_o \) are the position of vehicle and position of the obstacle in frame \( I \) respectively.

The LOS can be computed from the image feature of an object. Assuming zero angle of attack and zero side-slip angle, recalling (2) we can compute

\[
\begin{bmatrix}
    l_{lon} \\
    l_{lat}
\end{bmatrix} = \begin{bmatrix}
    \arctan \left( \frac{y_f}{x_f} \right) \\
    \arctan \left( \frac{y_f}{z_f} \right)
\end{bmatrix} \quad (20)
\]

where \( l_{lon} \) and \( l_{lat} \) denote the longitudinal and lateral LOS, respectively.

\section{III. IMAGE-BASED VISUAL SERVOING WITH NMPC FOR OBSTACLE AVOIDANCE}

\subsection{A. Problem definition}

The existing obstacle avoidance approaches have been studied about generating a set of waypoints \( q(1), \ldots, q(n) \) that allows the vehicle to avoid other objects. However, in practice, dynamic constraints (e.g., input/output constraint) can prohibit the vehicle from following the reference trajectory \( x_r(t) \) that passes waypoint \( q(k) \) at specific time \( t_k \). Therefore the reference trajectories have to be generated based on the vehicle dynamics, \( f_v \). It can be formulated as follows for some control input \( u(t) \).

\[
\begin{align*}
(\text{Solve}) & \quad x_r(t) \quad (9) \\
\text{such that} & \quad x_r(t_k) = q(k) \quad k = 1, \ldots, n \quad (10) \\
\text{and} & \quad \dot{x}_r(t) = f_v(x_r(t), u(t)) \quad (11)
\end{align*}
\]

Fig. 5. Geometry and coordinate frames for a pinhole camera model

where

\[
 J_p = \begin{bmatrix}
 -f & 0 & \frac{U_s V_c}{f} & -\frac{f^2 + U_s^2 V_c^2}{f} & V_s \\
 0 & -f & \frac{U_c V_c}{f} & -\frac{f^2 + U_c^2 V_c^2}{f} & -U_s \\
 \end{bmatrix} \quad (6)
\]

is called the image Jacobian matrix \( J_p \) associated with \( p \). (5) describes the relationship between the camera velocity defined in the camera frame and the image feature velocity defined in the image frame.

\subsection{C. Image and UAV Dynamics Integration}

Let the camera be mounted on the nose of the UAV looking forward. Therefore, the translational motion of the camera along with the \( z \)-axis \( T_z \), defined in the camera frame, is the same as the velocity of the UAV, \( V \). \( T_x \) and \( T_y \) are considered as zero with the assumption of no sideslip motion. Also, there is no rolling motion in the UAV dynamics as given in (1), so \( \omega_z \) is zero. The translation and rotational velocity of the camera \( \dot{r} \) can be regarded as

\[
[T_x, T_y, T_z, \omega_x, \omega_y, \omega_z] = [0, 0, V, \dot{\psi}, \dot{\gamma}, 0] \quad (7)
\]

Then, using a combination of (1) and (7), (5) can be written as (8):

\[
\dot{p} = \begin{bmatrix}
 \frac{U_c V_c}{f} + \frac{U_s V_s}{f} \cdot \frac{u_v}{V} - \frac{f^2 + U_s^2 V_s^2}{f} \cdot \frac{u_v - g \cos \gamma}{V} \\
 -\frac{V_c}{f} \cdot \frac{u_v}{V} + \frac{U_c U_s}{f} \cdot \frac{u_v}{V} - \frac{f^2 + U_c^2}{f} \cdot \frac{u_v - g \cos \gamma}{V} \\
 \end{bmatrix}
 = f_v(p, u, \gamma) \quad (8)
\]

\section{III. IMAGE-BASED VISUAL SERVOING WITH NMPC FOR OBSTACLE AVOIDANCE}

\subsection{A. Problem definition}

The existing obstacle avoidance approaches have been studied about generating a set of waypoints \( q(1), \ldots, q(n) \) that allows the vehicle to avoid other objects. However, in practice, dynamic constraints (e.g., input/output constraint) can prohibit the vehicle from following the reference trajectory \( x_r(t) \) that passes waypoint \( q(k) \) at specific time \( t_k \). Therefore the reference trajectories have to be generated based on the vehicle dynamics, \( f_v \). It can be formulated as follows for some control input \( u(t) \).

\[
(\text{Solve}) \quad x_r(t) \quad (9)
\]

such that  
\[
x_r(t_k) = q(k) \quad k = 1, \ldots, n \quad (10)
\]

and  
\[
\dot{x}_r(t) = f_v(x_r(t), u(t)) \quad (11)
\]
$x_k$. From the control input and output sequences, an optimal control input $u^*$ is optimized by minimizing the following cost function:

$$\{u^*_k\}_{t=1}^{t+N} = \text{arg} \min_{\{u_k\}_{t=1}^{t+N}} J(\{x_k\}_{k=t}^{t+N}, \{y_k\}_{k=t}^{t+N}, \{u_k\}_{k=t}^{t+N})$$

(22)

From the optimal input sequence $\{u_k\}_{k=1}^{t+N}$, the optimal input $\{u_k\}_{k=t}$ is applied to the given system. Then a next optimal control input is computed using a next time horizon $\{t+1, \ldots, t+N+1\}$ (See Fig.1).

C. Constraints embedding

The goal of this research is to control the UAV for tracking the reference LOS in dynamic unknown environment while keeping the following requirements.

- Control input should not exceed $u_{max}$.
- During the avoidance maneuvering, the destination image feature should remain in certain LOS, $l_{sat}$ (visibility maintenance).
- The LOS of each obstacle should larger than $l_{th}$ when the distance between the UAV and the obstacle is smaller than $z_{th}$.

In order to generate trajectories to satisfy the above requirements, an additional cost term must be added. Note that the obstacles are detected by the camera and range sensors, LOS and depth information are used to define the obstacle cost.

$$G_k^N(u_k) = \begin{cases} \frac{1}{2}(|u_k| - u_{sat})^2, & \text{if } |u_k| - u_{sat} > 0 \\ 0, & \text{else} \end{cases}$$

$$G_k^{\text{los}} = \frac{1}{2}(|l| - l_{sat})^2, \text{if } |l| - l_{sat} > 0$$

$$J_{\text{obs}} = \sum_{j=1}^{m} \frac{\mu_{\text{obs}}}{(l_j^0 + \epsilon)(z_j^0 + \epsilon)} G_{k,j}^{\text{obs}}$$

(25)

where $l_j$ is a LOS of the $j$-th obstacle, $\epsilon \ll 1$ is adopted to avoid singularity, $z_j^0$ is the positive depth value from the UAV to $j$-th obstacle and $G_{k,j}^{\text{obs}}$ is

$$G_{k,j}^{\text{obs}} = \begin{cases} 1, & \text{if } |l_j^0| < l_{th} \text{ and } 0 < z_j^0 < z_{th} \\ 0, & \text{else} \end{cases}$$

(26)

D. Optimization process

The optimal control input $\{u^*(t)\}_{t=1}^{t+N}$ is computed from the following optimization in the NMPC setting [11].

$$\text{arg} \min_{\{u_k\}_{k=1}^{t+N}} J(\{x_k\}_{k=t}^{t+N}, \{y_k\}_{k=t}^{t+N}, \{u_k\}_{k=t}^{t+N})$$

(subject to) $x_{k+1} = f_d(x_k, u_k)$

(27)

where

$$\phi_N = \frac{1}{2} \ddot{x}_N P \ddot{x}_N$$

(28)

$$L(x_k, u_k) = \frac{1}{2} \dot{x}_k^T Q \dot{x}_k + \frac{1}{2} u_k^T R \ddot{u}_k$$

(29)

$\ddot{x} \triangleq x_d - x$, and $P$ and $Q$ are constant positive definite weighting matrices to penalize the deviation from the reference LOS. $R$ is a constant positive definite weighting matrix to penalize the control input magnitude.

The above cost function can be rewritten as

$$J_a = \phi_N + \sum_{k=0}^{N-1} \sum_{k=0}^{N-1} L(x_k, u_k) + \lambda_{k+1}^T [f_d(x_k, u_k) - x_{k+1}] + G_k$$

(30)

where $G_k = \mu_k G_k^N + \mu_{\text{los}} G_k^{\text{los}} + \mu_{\text{obs}} \sum_{j=1}^{m} l_{obs,j}^0$. Here, $\mu_{\text{obs}}, \mu_{\text{los}}$, and $\mu_{\text{obs}}$ denote weighting parameters for control input, LOS and obstacle threat, respectively, and $m$ is a number of obstacles.

Let the Hamiltonian function be

$$H_k = L(x_k, u_k) + \lambda_{k+1}^T f_d(x_k, u_k) + G_k.$$ 

(31)

Since we want to choose $\{u^*(t)\}_{t=1}^{t+N}$ that minimizes $J_a$, we take a look at

$$\frac{\partial J_a}{\partial (x_k, u_k)} = (\frac{\partial \phi}{\partial x_N} - \lambda_k^N) + \frac{\partial H_0}{\partial x_k} + \frac{\partial H_0}{\partial u_k}$$

$$+ \sum_{k=1}^{N-1} \left( (\frac{\partial H_k}{\partial x_k} - \lambda_k^T) + \frac{\partial H_k}{\partial u_k} \right)$$

(32)

By choosing the Lagrange multiplier vector as below to make the derivative of $J_a$ zero,

$$\lambda_N^T = \frac{\partial \phi}{\partial x_N}$$

(33)

$$\lambda_k^T = \frac{\partial H_k}{\partial x_k}$$

(34)

(32) becomes

$$\frac{\partial J_a}{\partial (x_k, u_k)} = \sum_{k=0}^{N-1} \left( \frac{\partial H_k}{\partial u_k} \right) + \lambda_0^T$$

(35)

The gradient of the Hamiltonian with respect to control inputs is

$$\frac{\partial H_k}{\partial u_k} = u_k^T R + \lambda_k^T f_d(x_k, u_k) + \frac{\partial G_k}{\partial u_k}$$

(36)

Then the optimal control $u^*$ can be computed using gradient descent optimization [11].

IV. RESULTS

Guidance systems that utilize LOS information are widely employed for fixed-wing airplanes. Examples include VHF omni-directional range (VOR) guidance systems at airports, and net landing systems for automatic recovery of UAVs [10]. During the reference LOS tracking operation, the UAV may be required to perform an autonomous obstacle avoidance maneuver. In this section, reference LOS tracking performance while avoiding static and moving obstacles and maintaining visibility is evaluated using vision-integrated camera-in-the-loop simulation (CILS) as shown in Fig. 6. In the CILS setup, the image features of the destination and obstacles are displayed based on their locations and the state variables of UAV. The camera obtains visual information
from the screen to compute LOS of each object. Using LOS and other state variables, the NMPC controller optimizes a control input. The performance of the proposed algorithm is validated compared with classic visual information-based obstacle avoidance algorithm. The classic approach uses PID control for obstacle avoidance instead of NMPC.

The mission of the UAV is to follow the reference LOS derived from the image feature of the destination. As depicted in Fig. 7, the initial position and attitude of the UAV are \([x_0, y_0, z_0, \phi_0, \theta_0, \psi_0] = [-600, -40, 70, 0, 0, 0]\). The position of the destination is \([x_{des}, y_{des}, z_{des}] = [0, 0, 100]\). The static obstacle is located at \([-450, -30, 77]\), and the dynamic obstacle is moving from \([-50, 225, 80]\) with the velocity of \([-3.6, -4.4, 0]\). The units of position, attitude and velocity state variables are meter, degree and meter/second, respectively. Thick and thin solid lines represent UAV trajectories using NMPC and PID, respectively.

Table II presents the values of the parameters used in the numerical simulation. If an object other than the destination is placed within the threshold values of LOS and depth \(Z\), the object is considered as an obstacle. As shown in Fig. 8, collision with the static and moving obstacles is avoided around \([-500, -30, 75]\) and \([-300, -20, 80]\) for both controllers (NMPC and PID). However, the NMPC controller generates more efficient trajectory for obstacle avoidance maneuver as given in Fig. 10.

Using the constraints in the NMPC framework, the image feature of the destination can be maintained in the valid area (±15 deg), while the image feature leaves the valid area in the PID framework as shown in Fig. 9. Also, the NMPC controller generates the control inputs for longitudinal and lateral acceleration within their saturation values, while the excessive control inputs of the PID controller are cut off at the saturation value. Overall, the NMPC controller uses more efficient input than the PID controller while satisfying the visibility constraints. The total magnitude of control input for the NMPC controller is 53.34% that of the PID controller.

V. Conclusions

This paper describes an image-based obstacle avoidance algorithm using NMPC framework for a fixed-wing UAV. With a camera and range sensor, the proposed algorithm generates an optimal input to avoid the collision with static and dynamic obstacles while keeping the visibility of the destination image feature. Constraints for the visibility, UAV dynamics and the input saturation are considered in the NMPC framework. In order to implement the image-based

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling Time, ((\Delta \tau))</td>
<td>0.05 s</td>
</tr>
<tr>
<td>Horizon Length ((N))</td>
<td>20</td>
</tr>
<tr>
<td>Focal Length of The Camera ((f))</td>
<td>695</td>
</tr>
<tr>
<td>Gravity ((g))</td>
<td>9.81 m/s²</td>
</tr>
<tr>
<td>Velocity of the UAV ((V))</td>
<td>5 m/s</td>
</tr>
<tr>
<td>Valid Control Inputs ((u_0))</td>
<td>(-0.2 &lt; u_0 &lt; 0.2) m/s²</td>
</tr>
<tr>
<td>Valid Control Inputs ((u_{sat}))</td>
<td>9.61 m/s²</td>
</tr>
<tr>
<td>Saturation of LOS ((l_{sat}))</td>
<td>±10 deg</td>
</tr>
<tr>
<td>Threshold of LOS ((l_{th}))</td>
<td>±20 deg</td>
</tr>
<tr>
<td>Threshold of Depth ((z_{th}))</td>
<td>200 m</td>
</tr>
</tbody>
</table>
obstacle avoidance using NMPC, UAV and image dynamics are integrated so that the proposed algorithm successfully predicts the image states for a receding horizon. Vision-integrated HILS is performed to validate the presented algorithm. Performance requirements are satisfied with the given constraints in the unknown dynamic environment.

VI. ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program Through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-003656).

REFERENCES