Abstract: This paper discusses how the texture of a spherical object can be unwrapped from several images obtained from stereo cameras. This is specifically aimed at applications such as fruit grading where spherical fruit travel on a conveyor while cameras capture images at regular intervals. Since this is a real-time application, the proposed algorithm has to be acceptably efficient, while ensuring that the whole surface of the fruit is accounted for in the selected areas of the images with minimum redundancy.

Key–Words: Spherical Model, Texture Extraction, Fruit Sorting Applications

1 Introduction

In recent years, the growth of technology and the increasing demand for quality of consumer products, have led to the use of automated visual inspection systems in product lines. Although humans are endowed with the ease of detection of even a slight variation in the product due to the combination of senses such as sight, smell and taste, their performance varies under different conditions. Fatigue and boredom are two such factors and may cause inconsistency in the detection of faults.

The aim of an automated system is to remove these human inconsistencies and introduce a fast and consistent way of identifying defects. Examples of such automated systems are found in [2, 5, 8]. Most of these systems are specialized for an application and do not offer universal inspection of any object. The focus of this paper is on spherical objects, specifically fruit that travel forward on a conveyor. Hence, our method is also application specific.

Work such as [1, 6, 9] discuss different methods of automatically inspecting fruit on conveyors. An interesting point in most of these algorithms is that they either use the whole image, or a strip from the middle of each image for the feature detection to be used in blemish identification. When taken over several views, this introduces redundancy in data and causes unnecessary processing. The method discussed in this paper focuses on removing this redundancy and selecting a unique area from each of the images at the cost of some extra calculations in the texture extraction stage. The unwrapped texture can then be used for blemish detection or colour classification of fruit.

The rest of the paper is organized as follows: Section 2 introduces the theory, while section 3 discusses how the fruit is reconstructed as a sphere. Section 4 focuses on the selection of a unique area on the sphere viewed by each of the cameras in the relevant view. In section 5, we calculate the projection of the selected area on the images, while in section 6, we discuss experimental results.

2 Review - Quadrics & Projections

A quadric is a surface in 3d, represented by a $4 \times 4$ symmetric matrix $Q$. For a point $X$ on $Q$, given in homogeneous coordinates $X = [x \ y \ z \ 1]^T$, the following condition should be satisfied.

$$X^TQX = 0 \quad (1)$$

A sphere is a special case of quadrics, and is represented by a $4 \times 4$ symmetric matrix $Q_s$ as shown in eqn (2). For a point on the sphere, $X$, the equation $X^TQ_sX = 0$ is satisfied.

$$Q_s = \begin{bmatrix} I & -c \\ -c^T & \gamma \end{bmatrix} \quad (2)$$

where, $c$ is the center of the sphere, and $r^2 = c^Tc - \gamma$ is the square of the radius.
A conic is defined in 2d, similar to a quadric by a 3 × 3 symmetric matrix, \( C \). For a point \( x \), represented by a homogeneous vector of three elements lying on the conic, the constraint \( x^T C x = 0 \) is satisfied.

Duality in 3d is expressed as the interchangeability of points and planes [4]. Considering the equation of a plane, \( \pi^T X = 0 \), we see that interchanging the four element vectors would not change the equation. Hence, by substituting the polar plane \( \pi \) of the point \( X \) with respect to the quadric \( Q \) in eqn 1, we get:

\[
\pi^T Q^* \pi = 0
\]  

(3)

where, \( \pi = Q X \) and \( Q^* \) is the dual of \( Q \), defined by a set of planes and is obtained from \( Q^* = Q^{-1} \) or \( Q^* = \text{adj}(Q) \), whichever exists.

The outline of the projection of a quadric (ocluding contour) is formed by the contour generator, which is the curve formed by the tangency between the quadric and the cone of tangent light rays having the camera center as its apex. Since the contour generator arises from tangency, it is more convenient to use the dual quadric of \( Q \) in the determination of the projected conic. Under the camera matrix \( P \), the projection of \( Q \) is the conic \( C \) given by eqn (4). Here, \( Q^* \) and \( C^* \) are the duals of the quadric and the conic respectively. Proof of this is given in [4].

\[
C^* = PQ^*P^T
\]  

(4)

3 Reconstruction of Spheres

In our application, we have two cameras placed on either side of the conveyor with rollers that rotate the fruit as they move forward. The cameras are synchronized to get two shots of the fruit in the same position. Each camera captures three fruit in an image. Hence, we get six views of a fruit from different positions.

First, we create an approximate model of the fruit as a sphere using the stereo images at each position. Spherical reconstruction is a subproblem of quadric reconstruction discussed in [12]. The sphere specific method given in [11] was found to be more efficient and accurate for spherical objects and hence is used in this application.

We first fit ellipses to the segmented images of the fruit using the algorithm introduced in [10]. Then, the ellipses are adjusted to fit epipolar tangent constraints using the method presented in [12]. Since the cameras are calibrated, we know their intrinsic and extrinsic parameters using which, we calculate the cones tangent to the object having the camera centers as the apices. It was shown in [11] that the axis of the tangent cone goes through the center of the sphere and hence, it is uniquely defined by the two tangent cones obtained.

Then, to find the radius of the sphere, a frontier point [7] is used. It is defined as the fixed point on the surface of the quadric, corresponding to the intersection of two corresponding contour generators. Since frontier points lie on the surface, they satisfy eqn (1) giving us a unique sphere.

We model three spheres for the three different positions of the fruit using this method and determine the mean model to represent it. The surface of the model is then split into different areas to be viewed in each of the images.

4 Unique Area Determination

The axis of rotation of the fruit is along the direction connecting the two cameras and they move forward parallel to the ground in the direction normal to the axis of rotation. These two directions along with the normal to the ground plane, form a right-handed coordinate system. Images are taken every \( t_s \) seconds and the rotation angle (\( \theta \)) and the forward movement (\( d \)) during this interval is known.

To develop the theory, we assume that the rotational movement of the fruit is slip-less, and that the forward movement is significantly small. The latter assumption is valid in our system because the distance between views is very small. If not, we need to incorporate the effect of the forward movement into the rotation, and the selection of the areas would be more complex.

First, to find the areas viewed by each camera, we divide the spherical model into two hemispheres. Each hemisphere is best viewed by the camera situated on the same side as itself. Then, we divide each hemisphere into a number of areas to be selected by each positional view. This division is done through the contact point of the model and the axis of rotation (\( P_f \)). Figure 1 illustrates this.

To determine the angle that should be covered by each image, we need to consider several factors. Firstly, if the total angle that is covered by each view is \( \alpha \), the number of views that are available is \( n_v \), and the angle of the area coming into view at each rotation (rotation angle) is \( \theta \), eqn (5) should be satisfied to cover the whole surface.
\[ \alpha + (n_v - 1)\theta \geq 2\pi \implies n_v \geq \frac{2\pi - \alpha}{\theta} + 1 \quad (5) \]

Hence, we select the smallest integer that satisfies eqn (5) as the number of views to be used \( n_u \). If \( n_u \) is greater than the number of available views \( n_v \), it is not possible to capture the whole surface. If it is lesser than the available views, we select \( n_u \) of the centermost views, as the closer the object is to the camera, the better the projection of the surface. Secondly, \( \alpha \geq \theta \) should be satisfied, so that the rotation does not make an angle of more than \( \alpha \) come into view, thereby missing parts of the surface.

The next step is to determine the angle of the area that should be covered by each view. The simplest method of division is to get an angle of \( \beta \), related to the number of selected views \( n_u \), as follows.

\[ \beta = \frac{2\pi}{n_u} \quad \text{and} \quad \beta \leq \theta \quad (6) \]

For our application, the rotation angle \( \theta \) is close to \( \frac{2\pi}{3} \) and hence, we need three views of the fruit and get the complete texture by dividing it into three areas.

Before the division of areas, we need to find the middle of the centermost view so that the area selected could be centered around it. If there are two center views, we select one arbitrarily. We define this as the intersection point of the sphere and the line connecting the camera center and the center of the sphere. The solution of the sphere and the line gives a quadratic, from which we get two possible solutions. Only one lies between the camera center and the center of the sphere and is selected as the center of view \( P_c \).

From this, we find the center plane \( (\pi_c) \) for the centermost image. The plane \( \pi_c \) goes through the rotation point, center point \( P_c \), and the center of the sphere. The area for the center view is defined by boundaries \( \frac{\beta}{2} \) angle away from either side of \( P_c \). These are calculated by rotating the center plane \( \pi_c \) around the axis of rotation by the relevant angle and then finding the intersection between the sphere and the boundary planes. The boundaries are essentially circles with a center and radius the same as that of the sphere because it is formed as the intersection of a plane through the center of the sphere. The other boundaries are found similarly by changing the rotation angle. For example the next boundary is \( \frac{3\beta}{2} \) away from the center plane.

Then we assign these areas to the positional views. The area containing \( P_c \) is assigned to the center position and those before that are the areas coming into view before the center view. For example, area 2 in figure 1 is for the center view, while areas 1 and 3 are for the views before and after it respectively. These boundaries are still represented with respect to the centermost view and have to be rotated by \( n_r\theta \) and translated by \( n_rd \) to get the actual equations for the other views. Here \( n_r \) is the number associated with the view.

### 5 Projection of the Area

#### 5.1 Projection of the Boundaries

The projection of the circles (a special case of general conics) which form the boundaries of the areas on the image planes has to be determined next. For this, we use the duality property in two and three dimensions discussed in section 2.

First, we represent a conic in 3d by a plane \( (\gamma) \) and a 2d curve \( (C_p) \) on that plane. The plane is represented...
by a four element vector and imposes three degrees of freedom (dof). The 2d curve (conic) is given by a 3 x 3 matrix and has five dof. Hence, the plane conic has eight dof in all.

A conic in 3d is formed by the intersection between a cone $Q_c$ and the plane $\gamma$. See figure 2. The cone $Q_c$ is a degenerate quadric with eight dof and its dual is a plane conic $Q^*_c$ [3]. Conversely, the dual of a plane conic is a cone. As shown in [3], the dual of the plane conic $C_p$ is the cone passing through $Q^*_c$ (the dual of $Q_c$, which is a plane conic) and has the dual of the plane $\gamma$ (which is a point) as its apex.

Next, we define a transformation that converts the 2d coordinates of points on the plane $\gamma$ to 3d. For this, we create a 3d coordinate system from the 2d system on $\gamma$ with respect to which the conic $C_p$ is defined. The $x$ and $y$ axes and the origin are the same as that of the 2d system while the $z$ axis is in the direction of the normal to the plane, forming a right-handed system. Hence, any point on $\gamma$ can be written with respect to the new coordinate system, simply by giving the $z$ coordinate a value of zero. Then, we find the transformation between the world and plane coordinates as a rigid body motion.

Let this rigid body transformation that converts a point $X_p$ in 3d plane coordinates, to world coordinates $X$ be given by a $4 \times 4$ matrix $g$ which is known (since $\gamma$ is known). Then, the relationship, $X = g X_p$ holds. Note that the points are given in homogeneous coordinates. But, any point represented by $X_p$ is planar and therefore can be given as: $X_p = [x \ y \ 0 \ 1]^T$. If we define a $4 \times 3$ matrix $M$, by removing the third column of $g$, and a point on $\gamma$ in 2d plane coordinates $x = [x \ y \ 1]^T$, we get eqn (7).

$$X = M x$$

In the 2d plane coordinate system, $C_p$ satisfies the equation, $x^T C_p x = 0$. If the dual of $C_p$ on the plane is $C^*_p$, it is defined by the envelope of lines $1$ as follows:

$$1^T C_p^* 1 = 0$$

The dual of the plane conic in 3d $C^*_p$, which is a cone as described above, is defined by a set of planes (given in homogeneous coordinates), $\pi$.

$$\pi^T C_p^* \pi = 0$$

From $\pi^T X = 0$ and eqn (7), we get $\pi^T M x = 0$. This and $1^T x = 0$ gives us $1 = M^T \pi$. Substituting this in eqn (8), we get:

$$\pi^T M C^*_p M^T \pi = 0$$

Now, by comparing eqns (9) and (10), we get the dual of $C_p$ in 3d as follows:

$$C^*_p = M C^*_p M^T$$

$C^*_p$ (which represents the dual of a boundary curve) is then used in eqn (4), to find the projection on the image plane. The dual of the projected conic $C^*_i$ is obtained here, the inverse/adjoint of which gives the actual projection $C_i$.

$$C^*_i = PC_p M^T$$

### 5.2 Selection of the Area

By projecting the boundaries of the areas (plane conics), we get the 2d area that should be extracted from each image. The projections, along with the selected area for a spherical object for an image from one of the stereo cameras, is shown in figure 3. How the selection of this area is done from the conic boundaries, is shown in figure 4. It is divided into four areas as shown and their union is taken to form the composite area which gives a unique part of the surface of the fruit in each image.
Let $C_1$, $C_2$ and $C_3$ be the projected boundaries and $P_1$, $P_2$ and $P_3$ be their intersection points. To determine the lines that form the triangle $P_1P_2P_3$, the intersection points of the conics are determined. For this, solving of pairs of quadratic equations of two variables is not required, as a simpler method of calculation is possible. We calculate the intersection points of the boundaries in 3d and project them to the image planes to get their 2d counterparts.

Two of the boundaries always intersect at the rotation point, as seen in figure 1. The other two points are at the intersection between the hemispheric boundary and one of the other two boundaries. This is easily calculated as it lies on the intersection of the two planes containing the boundaries, at a distance of $r$ (radius of the model) away from the center of the model. Projecting them to the image plane, we calculate the lines $P_1P_2$, $P_2P_3$, and $P_3P_1$ that pass through them.

Then, we select areas according to these lines and the conics. For example, $Area\ 1$ is bound by the three lines and the other areas lie between the lines and the conics, as shown in figure 4. These calculations are simple and involve the substitution of points to the conic/line equations and checking for the sign of the result. After the selection of the separate areas are done, the union is taken to get the selected area.

## 6 Experimental Results

The algorithm discussed above was applied in the target system. For the experiments, spherical objects with different amounts of blemishes on them were used. Six views each were obtained from the two cameras at three positions and the texture was unwrapped.

A comparison was done with some methods of texture extraction used in the fruit grading industry. Figure (5) shows these different methods. The first (named Full) is the inspection of the full object extracted from the background. The next method is the most common way of texture inspection, and is done by considering a strip of surface from the center of the image [1]. We use two such instances with the width of the strip being 75% and 50% of the width of the object as captured in the image. These are referenced as Strip 75% and Strip 50% respectively. Finally, we introduce our algorithm of texture extraction (Opt. Area) into the comparison.

The comparison is done for the following categories: missed blemishes $M$, duplicate detections $D$ and partial detections $P$. If the number of blemishes is $n$, the total number of detected blemishes is $t$, the number of missed blemishes is $m$, and the number partially shown is $p$, we define the quantities of error as follows:

\[
M = \frac{m}{n} \times 100\% \quad (13)
\]
\[
D = \frac{(t - n)}{n} \times 100\% \quad (14)
\]
\[
P = \frac{p}{t} \times 100\% \quad (15)
\]

The average values of these errors were calculated for all four methods. Table 1 summarizes the results.

From the experimental results we observe that our method has very low levels of error with respect to all compared categories. The missed and partial detections were due to the slipping of the objects on the rollers as they rotate on the conveyor. This makes the angle of rotation different from the calculated one and introduces the observed errors. Considering a slightly larger window for point selection rectifies this problem but increases the duplicate detections. Hence, the adjusting of the size of the window should be done according to the expected outcome.

With respect to speed, our method was found to be 4 times slower than the method using the full image and 1.4 times slower than the one selecting a strip from the center of the image. This is due to the additional calculations required for creating the model and projecting the boundaries. The modelling of the fruit (as ellipsoids or spheres) is a necessary preliminary step in our target application as it is required for size sorting. Hence, the additional processing (when compared to the other methods described) is justified and hence is acceptable.

### Table 1: Experimental Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Missed %</th>
<th>Duplicate %</th>
<th>Partial %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>0.00</td>
<td>153.33</td>
<td>28.95</td>
</tr>
<tr>
<td>Strip 75%</td>
<td>3.33</td>
<td>106.67</td>
<td>36.84</td>
</tr>
<tr>
<td>Strip 50%</td>
<td>43.33</td>
<td>3.33</td>
<td>17.11</td>
</tr>
<tr>
<td>Opt. Area</td>
<td>3.33</td>
<td>0.00</td>
<td>5.26</td>
</tr>
</tbody>
</table>

## 7 Conclusion

We discussed an algorithm to unwrap the texture of a spherical object given several views from stereo cameras. It ensures that an acceptable number of pixels are selected and that the redundancy of selected pixels is minimized. This reduces calculations in the next stage.
(blemish detection) but is slower than the other methods used for surface extraction. This compromise is acceptable because most of the additional calculations are related to the preliminary step of modelling which is also required for the size sorting of fruit.

This method has the additional advantage that selecting a unique area from each of the images makes it easier to map the texture back to the model to view the whole surface in 3d. This can be used in colour sorting of fruit as well, with or without the mapping to the 3d model.

Future work involves the extension of the algorithm to suit ellipsoidal models as a broader range of fruit can be categorized as ellipsoidal. Further, modification should be done to account for the slip in the rotational movement to increase the accuracy.

References:


