https://doi.org/10.2352/ISSN.2470-1173.2019.12.HVEI-209


# On the role of edge orientation in stereo vision

***Alfredo Restrepo and Julian Quiroga***
***Dpt. of Ingeniería Electrónica***
***Pontificia Universidad Javeriana; Bogotá, Colombia***

## Abstract

*As a biologically inspired guess, we consider two stereo information channels. One is the traditional channel that works on the basis of the horizontal disparity between the left and right projections of single points in the 3D scene; this channel carries information regarding the absolute depth of the point. The second channel works on the basis of the projections of pairs of points in the 3D scene and carries information regarding the relative depth of the points; equivalently, for a given azimuth disparity of the points, the channel carries information of the ratio of the orientations of the left and right projections of the line segment between the pair of points.*

## Introduction

In a previous work [5] we remarked on the conspicuous fact that edge orientation and ocular dominance are made simultaneously explicit in the visual cortex; here, we explore in particular the aspect of edge orientation and the role it may be playing in the computation of stereopsis.

The left halves of the left and the right retinae (where the right view field is projected by the lenses) are processed by the left thalamus and the left V1, and viceversa. The central fovea is processed by V1 neurons of both hemispheres that are connected via the corpus callosum. An edge point in the right visual field has projections to V1 neurons of the left hemisphere of three types: binocular of several degrees of ocularity, monocular driven by the left eye and monocular driven by the right eye. The topographic cortical distance between the monocular neurons is related to the horizontal disparity and we hypothesise that the firing binocular neurons lie in an topographic interpolation path of minimal orientation variation, between the corresponding monocular neurons.

The retinal output conveys information regarding edges; it is also known that, at cortical area V1 of old world monkeys, cells make explicit the attribute of orientation of edge segments, at various degrees of ocular dominance. From a single 3D scene two visual fields result, one for each retina; at V1, the two fields are interleaved via cells of different degrees of ocular dominance and, somehow, correspondingly, a single visual field is perceived, as if with a single, centred eye.

There is not a proven model for how the brain extracts the information about depth via stereopsis. It is known that the stereo channel belongs to the ”where system” [3] and it is also known that, at cortical area V1, many neurons are sensitive both to orientation and to ocular dominance. As we show here, for a small rectilinear edge segment that is specified by its two end points, the difference of orientations of the edge, as seen by the left and right eyes, is in a one-to-one correspondence with the difference of horizontal disparities corresponding to the two points, which in turn is in one-to-one correspondence with the difference of depths of the points. This is perhaps the reason why orientation and orientation disparity are important in human vision.

It may be convenient to think that to each pair of points in the 3D scene being watched there corresponds the virtual line segment that joins them. If none of the points is occluded, neither at the left nor at the right projection, the virtual line segment projects on the focal plane with orientations that, generally speaking, are different. This difference of orientations, or *orientation disparity*, depends both on the relative depth points and on the difference of their azimuths. We consider stereo algorithms that compute relative depth and orientation, as a biologically inspired model.

### *Terminology*

A few remarks regarding the terminology are perhaps in order. Being affine linearly related to disparity, we deal with the inverse of depth and call it *nearness* [6]. We use a single Cartesian coordinate system, instead of using one coordinate system for each of the two cameras; thus, the common focal plane is simply $z = \mathrm{f}$, where f is the focal distance [1]. We assume the camera pair to be centred at the origin of a unique coordinate system of points $[x, y, z]$. We assume that the $x$ coordinate is vertical and measures *height*, that the $y$ coordinate is horizontal, and the $z$ coordinate measures *depth*. By *distance* we mean the Euclidean distance $\sqrt{x^2+y^2+z^2}$ from the origin of the coordinate system. The term *orientation* is borrowed from the works of Hubel and Wiesel [2]; they placed lines and edges of different *orientations* on a vertical screen, as stimuli to neurons in cortical area V1. By *azimuth*, or planar direction, we mean the slope $\frac{z}{y}$ of the line from the origin to the orthographic projection $[0, y, z]$ of the point on the horizontal yz plane at zero height. When only one point is considered, the measures are said to be absolute; when two points are involved the measures are said to be relative.

## Absolute nearness

We derive first a well known formula, in order to set the notation for the remaining of the paper. In the single Cartesian coordinate framework $[x, y, z]$, $x$ is the vertical coordinate, $z$ is the depth coordinate and $y$ is the lateral horizontal coordinate. Points in the 3D scene are watched by a computer vision system equipped with two pinhole cameras, positioned at $[0, -A, 0]$ and $[0, +A, 0]$, in a parallel set up; the distance between the pinholes, or baseline, is $2A$. For a given point $[X, Y, Z]$ in the 3D scene consider the rays from the left and right pinholes to the point. For convenience, consider the projection, focal plane to be placed in front of the pupils, at $z = \mathrm{f}$, where f is the focal distance, rather than at the back of the pupils. The left and right radial projections of the point on the focal plane are found with respect to the left and right pupils; in barycentric coordinates, the *projecting lines* are

$$(1-\lambda)[0, \pm A, 0] + \lambda[X, Y, Z],\ \lambda \in \mathbf{R} \tag{1}$$

The projecting lines intersect the focal plane when $\lambda = \frac{f}{Z}$ and so the left and right projection points are

$$p_L = f[\frac{X}{Z}, \frac{Y+A}{Z} - \frac{A}{f}, 1], \; p_R = f[\frac{X}{Z}, \frac{Y-A}{Z} + \frac{A}{f}, 1]. \quad (2)$$

The only coordinate of the projections that is different is the horizontal, y-coordinate; the vertical coordinates, being on the equipolar plane, are the same. Calling the horizontal coordinates $Y_L$ and $Y_R$, you have thus

$$Y_L = f(\frac{Y}{Z} + \frac{A}{Z} - \frac{A}{f}), \text{and}, \; Y_R = f(\frac{Y}{Z} - \frac{A}{Z} + \frac{A}{f}). \quad (3)$$

Their difference is the (horizontal) disparity; calling it $\Delta_Y$, you have

$$\Delta_Y = Y_L - Y_R = 2f(\frac{A}{Z} - \frac{A}{f}). \quad (4)$$

The nearness $\zeta$ of the scene point to the origin of the coordinate system is then

$$\zeta = \frac{1}{Z} = \frac{1}{f}(\frac{\Delta_Y}{2A} + 1) \quad (5)$$

the nearness is thus *affine linearly* related[1] to the disparity, an appropriate feature if we think that nearer objects are more relevant to the vision system; therefore, when two points are considered (as we do below) their relative nearness is proportional to their disparity difference.

Also, from Equation 3, the average of $\frac{Y_L}{f}$ and $\frac{Y_R}{f}$ gives $\frac{Y}{Z}$:

$$\frac{Y}{Z} = \frac{1}{2}(\frac{Y_L}{f} + \frac{Y_R}{f}); \quad (6)$$

we call $\frac{Z}{Y}$ the *azimuth* of the point $[X,Y,Z]$. The azimuth is thus the slope of the line through the origin of the coronal yz plane, and the orthographic projection $[0,Y,Z]$ of the point $[X,Y,Z]$, on this plane. Equation 6 gives the (inverse) azimuth of the point, computed from the left and right projections of the point.

## Relative nearness of a pair of points in the 3D scene

In mammal vision, the retinal output makes explicit edge points, while the primary visual cortex makes explicit edge orientation at different degrees of ocular dominance. For each pair of points $q$ and $q'$ in the 3D scene, the virtual (or actual) line segment between them determines left and right projected line segments at the focal yx plane, which in general have have different slopes or orientations; see Figure 1.

### *The relative nearness linearly depends on the difference of disparities*

Given two points $q = [x,y,z]$ and $q' = [x',y',z']$, as in Figure 1, we compute their *relative nearness*. The left and right projections of the points are

$$q_L = f[\frac{x}{z}, \frac{y+A}{z} - \frac{A}{f}, 1], \quad q'_L = f[\frac{x'}{z'}, \frac{y'+A}{z'} - \frac{A}{f}, 1]$$

$$q_R = f[\frac{x}{z}, \frac{y-A}{z} + \frac{A}{f}, 1], \quad q'_R = f[\frac{x'}{z'}, \frac{y'-A}{z'} + \frac{A}{f}, 1].$$

Letting

$$\Delta := q_L - q_R, \; \Delta' := q'_L - q'_R$$

and

$$\Delta_L := q_L - q'_L, \; \Delta_R := q_R - q'_R$$

the *difference of disparities* i.e. the disparity corresponding to the point $q$ minus the disparity corresponding the point $q'$, denoted as $\partial_D$, is given by

$$\partial_D := \Delta - \Delta' = (q_L - q_R) - (q'_L - q'_R) \quad (7)$$

and note that it is also is also the difference of the differences, that for the left eye minus that for the right eye, of the projections of the two points:

$$\partial_D = (q_L - q'_L) - (q_R - q'_R) = \Delta_L - \Delta_R. \quad (8)$$

That is, calling $\Delta_L$ and $\Delta_R$ the *closeness* of the projected points at the focal plane, as seen separately by the left and right eyes, the difference of disparities is equal to the difference of closenesses.

We now derive a formula for the horizontal component of $\partial_D$; from

$$\Delta_L := q_L - q'_L = f(\frac{x}{z} - \frac{x'}{z'}, \frac{y+A}{z} - \frac{y'+A}{z'}, 0)$$

and

$$\Delta_R := q_R - q'_R = f(\frac{x}{z} - \frac{x'}{z'}, \frac{y-A}{z} - \frac{y'-A}{z'}, 0).$$

you get

$$\partial_D = \Delta_L - \Delta_R = 2Af(0, \frac{1}{z} - \frac{1}{z'}, 0) = 2Af(0, \zeta - \zeta', 0).$$

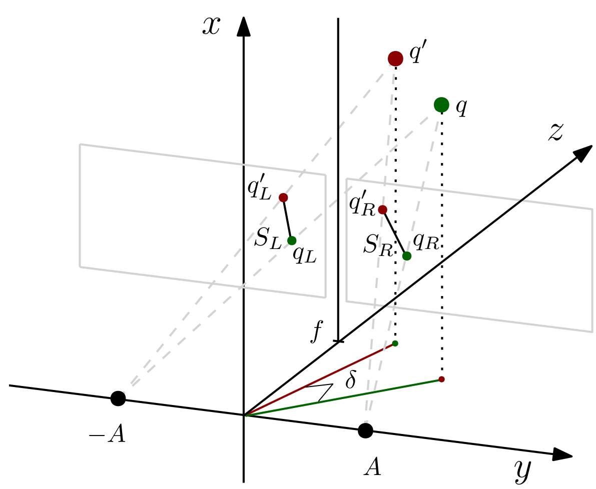



***Figure 1.*** *Projections of points $q$ and $q'$ on the focal plane. The point $q = [x,y,z]$ projects at $q_R$ with respect to the right camera, and at $q_L$ with respect to the left camera; likewise for the point $q' = [x',y',z']$. The points $q_R$ and $q'_R$ determine a line with slope $S_R$ in the projection plane, and the points $q_L$ and $q'_L$ determine a line with slope $S_L$. The azimuths of the points $q$ and $q'$ are given by the slopes $\frac{z}{y}$ and $\frac{z'}{y'}$; $\delta$ is the difference of azimuths.*

[1] $\zeta - \frac{1}{f}$ is *linearly* related to the horizontal disparity $\Delta_Y$.

The horizontal, $y$ component, of the difference of disparities is thus

$$\partial_{DY} = 2Af(\zeta - \zeta') \tag{9}$$

and the *relative nearness* is proportional to it:

$$\Delta_\zeta := \zeta - \zeta' = \frac{1}{2Af}\partial_{DY} \tag{10}$$

We now compute the ratio of the slopes, or *orientations*, of the projections, for the left and right eyes, of the line segment between a pair of points in 3-space, on the focal y-x plane $z = f$; this ratio depends as well on the relative nearness of the points.

### Orientation (slope) disparity

Consider the the slopes of the left and right projections of the line segment between the points $q = [x,y,z]$ and $q' = [x',y',z']$, on the focal y-x plane. We relate the slopes via their quotient and call it the *slope disparity* $\partial_S$ of the projections.

The barycentric equation of the line through the points $q$ and $q'$ is $(1-\lambda)q' + \lambda q$, or, component wise,

$$[(1-\lambda)x' + \lambda x, (1-\lambda)y' + \lambda y, (1-\lambda)z' + \lambda z].$$

The left and right projections on the focal plane, of this line through $q$ and $q'$, are given by

$$\Pi_L(\lambda) = f[\frac{(1-\lambda)x' + \lambda x}{(1-\lambda)z' + \lambda z}, \frac{Y + A}{(1-\lambda)z' + \lambda z} - \frac{A}{f}, 1] \tag{11}$$

and

$$\Pi_R(\lambda) = f[\frac{(1-\lambda)x' + \lambda x}{(1-\lambda)z' + \lambda z}, \frac{Y - A}{(1-\lambda)z' + \lambda z} + \frac{A}{f}, 1]. \tag{12}$$

To get the slopes of the projected lines consider the segment between the points $q'$, when $\lambda = 0$, and $q$, when $\lambda = 1$; the left and right projections of $q'$ are

$$\Pi_L(0) = f[\frac{x'}{z'}, \frac{y'+A}{z'} - \frac{A}{f}, 1], \text{ and, } \Pi_R(0) = f[\frac{x'}{z'}, \frac{y'-A}{z'} + \frac{A}{f}, 1]$$

while the left and right projections of $q$ are

$$\Pi_L(1) = f[\frac{x}{z}, \frac{y+A}{z} - \frac{A}{f}, 1], \text{ and, } \Pi_R(1) = f[\frac{x}{z}, \frac{y-A}{z} + \frac{A}{f}, 1].$$

In the focal y-x plane, as seen by the left and right cameras, the slopes $\frac{\Delta x}{\Delta y}$ are given by

$$S_L = \frac{f(\frac{x'}{z'} - \frac{x}{z})}{f(\frac{y'+A}{z'} - \frac{y+A}{z})}$$

$$= \frac{x'z - xz'}{(y'+A)z - (y+A)z'} = \frac{x'z - xz'}{y'z - yz' + A(z - z')} \tag{13}$$

and

$$S_R = \frac{f(\frac{x'}{z'} - \frac{x}{z})}{f(\frac{y'-A}{z'} - \frac{y-A}{z})}$$

$$= \frac{x'z - xz'}{(y'-A)z - (y-A)z'} = \frac{x'z - xz'}{y'z - yz' - A(z - z')}, \tag{14}$$

which are independent of $f$. We denote the *slope disparity* as

$$\partial_S := \frac{S_L}{S_R}, \tag{15}$$

which is independent of $x$ and $x'$; for the case of $y \neq y'$ (i.e. not on the same vertical line), you get

$$\partial_S = \frac{y'z - yz' - A(z - z')}{y'z - yz' + A(z - z')} \tag{16}$$

If $y = y'$ and $z \neq z'$ (points of different depth but on a same zx sagital plane)

$$\partial_S = \frac{S_L}{S_R} = \frac{\frac{x'z - xz'}{(y+A)(z-z')}}{\frac{x'z - xz'}{(y-A)(z-z')}} = \frac{y - A}{y + A} \tag{17}$$

unless $z = z'$, in which case the points $q$ and $q'$ lie on the same vertical line in 3-space and, at the focal plane, both projection segments are likewise vertical:

$$\Pi_L(0) = f[\frac{x'}{z}, \frac{y+A}{z} - \frac{A}{f}, 1], \ \Pi_R(0) = f[\frac{x'}{z}, \frac{y-A}{z} + \frac{A}{f}, 1],$$

$$\Pi_L(1) = f[\frac{x}{z}, \frac{y+A}{z} - \frac{A}{f}, 1], \text{ and, } \Pi_R(1) = f[\frac{x}{z}, \frac{y-A}{z} + \frac{A}{f}, 1];$$

that is, the coordinates $[y,x]$ of the projected points on the focal y-x plane, for the left pupil, are $f[\frac{y+A}{z} - A, \frac{x'}{z}]$ and $f[\frac{y+A}{z} - A, \frac{x}{z}]$, which determine a vertical segment of slope $S_L = \infty$; likewise, for the right pupil, the projection points are $f[\frac{y-A}{z} + A, \frac{x'}{z}]$ and $f[\frac{y-A}{z} + A, \frac{x}{z}]$, which also give $S_R = \infty$. Thus, the azimuths of $q$ and $q'$ are the same and their difference, which we call below $\delta$, in Equation 18, is zero.

If both points $q$ and $q'$ are of the same height $x = x'$, then the line through them projects on the focal y-x plane also in a horizontal fashion, and $S_L = S_R = 0$. Whenever the depth of the points is the same, i.e. $z = z'$, there is no slope disparity: $\partial_S = 1$.

### Relation between $\partial_S$ and $\partial_D$

The orthographic projections $[y,z]$ and $[y',z']$, on the coronal horizontal yz plane, of the points $q = [x,y,z]$ and $q' = [x',y',z']$ give the azimuths of the points., as

$$\alpha = \frac{z}{y}, \text{ and, } \alpha' = \frac{z'}{y'}$$

The inverses, $\frac{y}{z}$ and $\frac{y'}{z'}$, of the azimuth of the points $q$ and $q'$ are computed using Equation 6. The difference of the inverses of the azimuths is

$$\delta := \frac{y}{z} - \frac{y'}{z'}. \tag{18}$$

The difference $\delta$ of the inverses of the *azimuths* of the points $q$ and $q'$ is called the *azimuth disparity*. For $\partial_s \neq 1$, $\delta$ is a parameter that determines a family of 1-1 correspondences between the slope disparity $\partial_S$ and the difference $\partial_D$ of displacement disparities; see Equation 20 and Figure 2 below. In fact, given two values of $\partial_D$, $\partial_S$ or $\delta$, the third one is specified, except when $\partial_S = 1$ in which case $\partial_D = 0$, independently of the value of $\delta$; also, as explained below, $\partial_S = -1$ is a singular case.

From Equation 16, you get

$$\partial_S = 1 - \frac{2}{1 + \frac{y'z - yz'}{A(z - z')}} = 1 - \frac{2}{1 - \frac{2f}{\partial_D}\left(\frac{y'}{z'} - \frac{y}{z}\right)}$$

$$= 1 - \frac{2}{1 + \frac{2f}{\partial_D}\delta} \tag{19}$$

Unless $\partial_S = \pm 1$, for each point $(\partial_D, \partial_S)$, in the plane $\partial_S - \partial_D$, there corresponds a unique $\delta$ that makes the correspondence valid; see Figure 2. More concisely, you get

$$\partial_D = 2f\delta \frac{1 - \partial_S}{1 + \partial_S} \tag{20}$$

unless $\delta = 0$, in which case $\partial_S = -1$, independently of $\partial_D$; see Figure 2. This case $\frac{y}{z} = \frac{y'}{z'}$ of equal azimuth occurs when the points $q$ and $q'$ project on the coronal yz plane on a same line through the origin.

Actually $\partial_D = 0$ if and only if $\Delta_\zeta = 0$ if and only if $\partial_S = 1$, independently of $\delta$, which happens whenever the two points have the same depth $z = z'$. In other words, it is equivalent for two points $q$ and $q'$ to have the same depth, to determine the same horizontal disparities $q_L - q_R$ and $q'_L - q'_R$, to be at the same horizontal distance $q_L - q'_L = q_R - q'_R$ in both left and right projections at the focal plane or, for the line segment between the points, to determine equal slopes $S_L = S_R$, in the vertical, focal, frontal yx plane.

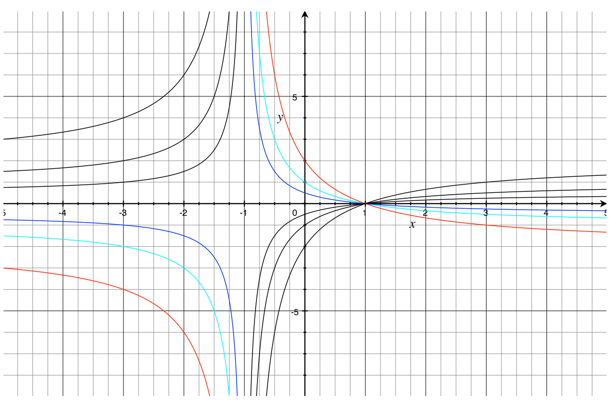


***Figure 2.*** *A plot of $\partial_D = 2f\delta\frac{1-\partial_S}{1+\partial_S}$ versus $\partial_S$, for $2f\delta = 0.5$ (blue), 1.0 (cyan) and 2.0 (red); also, for $2f\delta = -0.5, -1.0, -2.0$ (all in black).*

The value $\partial_S = -1$ is singular; see Figure 2 where, considered as a limit, the case $\partial_S = -1$ (of orthogonally projected segments), corresponds to $\delta = 0$ and an infinite value of $|\partial_D|$. For $\partial_S = 1$, as alreday said, for any $\delta$, you have $\partial_D = 0$[2].

### *Guidelines for an algorithm*

Classically, at each given height $x$ of the yx projection plane, for each pair of corresponding projection points $q_L$, $q_R$, the horizontal disparity is computed, from which you get the depth of the corresponding point $q$ in the 3D scene; then, you get its height as well. You may as well compute the average of the projections from which you get the azimuth of the point. You might call this the *absolute channel*.

By considering, as in V1, the left and right camera projections to be superimposed, one on top of the other, with the central or fixation points coinciding, the disparity becomes the distance between the projections of the point.

For a second, *relative channel*, instead of single single edge points, consider working with short edge segments, or equivalently, pairs of nearby edge points. In general, unless the segment is horizontal, the heights $x$ and $x'$ of the two extremes $q$ and $q'$ of the edge segment are different, and each of them will have left $q_L$, $q'_L$ and right $q_R$, $q'_R$ projections; note that the projections $q_L$ and $q_R$ are of the same height, and so are $q'_L$ and $q'_R$. From the differences of the projections you compute the nearnesses $\zeta$ and $\zeta'$, and from the averages of the projections you compute the azimuths $\alpha$ and $\alpha'$ of the points. From these data you compute now the difference of horizontal disparity $\partial_D = \Delta - \Delta'$, the azimuth disparity $\delta = \alpha - \alpha'$ and the orientation disparity $\partial_S = \frac{S_L}{S_R}$; for the correspondences to be valid, these data must comply with Equation 20. This channel tells you which object is nearer, without telling you the absolute nearnesses.

## Conclusion

The attribute of edge orientation is an early-vision descriptor that can be exploited by stereo algorithms. This can be done e.g. by matching pairs of projection points having the same heights, from the left and right projection images.

Stereo algorithms usually work at the pointwise level, matching either pixels or edge pixels; by considering pairs of pixels, the algorithm becomes *local* and short edge segments are matched.

The ratio of the orientation of projected 2D edge segments in the focal plane carries information that is equivalent to the relative depth of the end points of the corresponding 3D edge segment. Relative depth is a useful attribute in computer and robot vision; also, the use of the orientation of line segments, virtual or not, in the 3D scene being imaged has applications e.g. for robot surgeons.

For a given azimuth disparity, the orientation disparity and the difference of horizontal disparities give the same information. Thus, the use of the orientation disparity is not to be used to disambiguate stereo matches. It does provide however with an additional attribute that can generate additional matching criteria.

From the point of view of the architecture of V1, by superimposing at V1 the left and right retinal images, a rough correspondence results between horizontal disparity and cortical distance. Also, between cortical distance together with (occipital-temporal) laterality and azimuth. Orientation disparity should help to fuse the images.

The attribute of ocular dominance has not been exploited here; we are currently exploring this topic. We hypothesise that the attribute of ocular dominance is used by the cortex to find paths of minimal orientation variation, which it uses to solve correspondence ambiguities.

## References


[1] Benjamin T. Backus and Daniel Matza-Brown. The contribution of vergence change to the measurement of relative disparity. Journal of Vision. (2003)


[2] For points on a wire in 3-space, you may well have a zero crossing of $\partial_D$ as a function of wire length.

[2] Eye, Brain and Vision, http://hubel.med.harvard.edu/book/bcontex.htm
[3] Margaret Livingstone. Vision and Art. Abrams, New York, (2014).
[4] David Marr, Vision, Freeman, San Francisco (1982).
[5] A. Restrepo, Orientation-Ocularity maps: A technique for computer vision, IS&T International Symposium on Electronic Imaging - Human Vision and Electronic Imaging https://doi.org/10.2352/ISSN.2470-1173.2017.14.HVEI-132 (2017)
[6] J. R. Tresilian and M. Mon-Williams. Getting the measure of vergence weight in nearness perception. Experimental Brain Research, vol. 132, No. 3, pp. 362-368. (2000)

## Author Biography

*Alfredo Restrepo got his degree of Ingeniero Electrónico from the Pontificia Universidad Javeriana at Bogotá, and the M.Sc. and Ph.D. degrees from the University of Texas at Austin. He has made research stays at Oklahoma State University, the ENST in Paris and the University of Trieste. He was formerly at the Universidad de los Andes and is currently adjunct professor at the Departamento de Ingeniería Electrónica of the Pontificia Universidad Javeriana at Bogotá, Columbia.*

*Julian Quiroga got his degree of Electronics Engineer from Universidad Nacional de Colombia, the MSc degree in Electronics and Computers at the Universidad de los Andes in Bogotá and his PhD in Computer Vision at INRIA Grenoble, in France. He has been Associate Professor at Pontificia Universidad Javeriana in Bogotá. He is currently employed as Research Scientist at Genius Sports.*

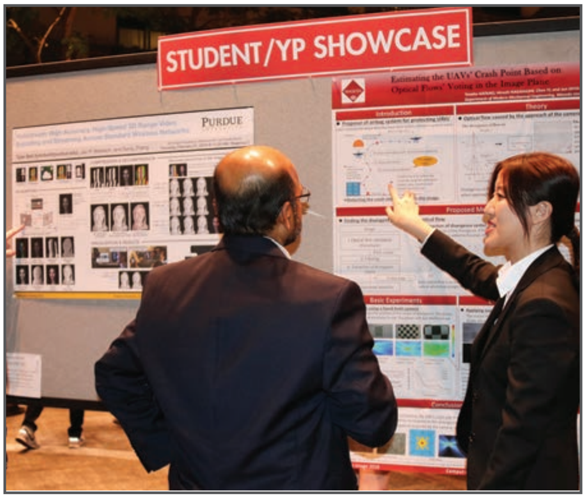
## JOIN US AT THE NEXT EI!

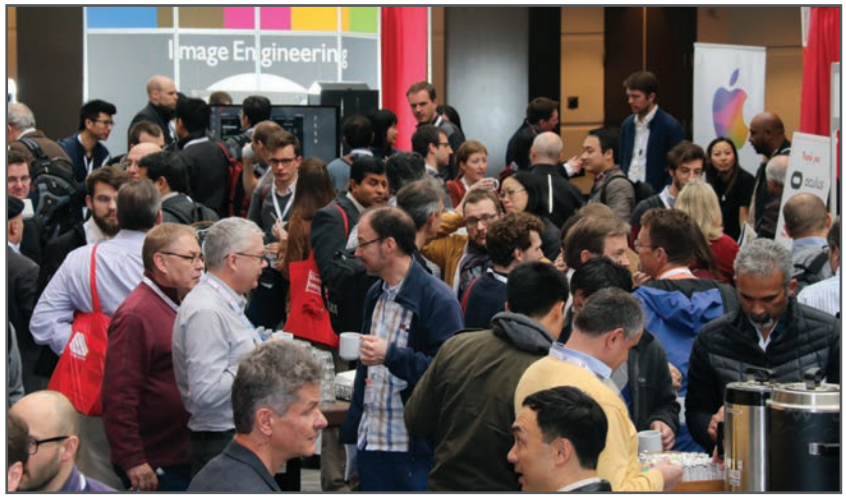
IS&T International Symposium on

# Electronic Imaging

SCIENCE AND TECHNOLOGY

*Imaging across applications . . . Where industry and academia meet!*




**• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •**

**• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •**

**www.electronicimaging.org**

IS&T
imaging.org