

OpenCV ve Kinect ile Stereo Görüntüden Derinlik Algılama

¹Hüseyin Kutlu ve ^{*2}Engin Avcı ve ^{*3}Ali İhsan Çelik

¹Besni Meslek Yüksekokulu, Bilgisayar Kullanımı Bölümü Adıyaman Üniversitesi, Türkiye

^{*2}Teknoloji Fakültesi, Yazılım Mühendisliği Bölümü Fırat Üniversitesi, Türkiye

^{*3}Besni Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü Adıyaman Üniversitesi, Türkiye

Özet

Bu çalışma bilgisayar görmesi alanında önemli bir yeri olan derinlik algılama hakkındadır. Çalışma kapsamında yapay görme alanında yaygın olarak kullanılan derinlik algılama tekniklerine değinilmiştir. Derinlik algılama tekniklerinden Stereovision ve benzer bir mantıkla çalışan Kinect (prime sence) ile derinlik algılama uygulamalarının nasıl gerçekleştirileceği üzerinde durulmuştur. Stereo Vision tekniği ile iki kameradan alınan görüntü açık kaynak kodlu görüntü işleme kütüphanesi OpenCV ile işlenmiş ve derinlik bilgisi edinilmiştir. Görüntü işleme ile derinlik algılamada bir diğer teknik olan Kinect (prime sence) ile derinlik algılama çalışması gerçekleştirilmiştir.

Anahtar Kelimeler: Derinlik algılama; Görüntü işleme ile derinlik algılama teknikleri; Stereo görüntü; Kinect

Abstract

This study is about depth detection which has an important role in the field of computer vision. Within this study we focused on the artificial vision field depth detection techniques commonly used. this study focuses on how to perform application of depth detection techniques which are stereo vision and Kinect (prime sence) that running a similar logic with stereo vision. Image taken from two cameras that treated with open source computer vision library (OpenCV) and depth information is acquired. Image taken from two cameras that treated with OpenCV and depth information is acquired.

Key words: Depth detecton, Depth detection tecniques with image processing, stereo vision, kinect

1. Giriş

İnsanların, birçok karmaşık davranışı, uzun pekiştirmeler olmadan gözleyerek, taklit ederek öğrendikleri görülmektedir. Söz konusu bilgisayarlara geldiğinde durum aynıdır. Bilgisayarların insan gibi düşünen, hesaplayan, çoğu donanımı insan uzuvlarına benzeyen, bir insan takliti olduğu söylenebilir. Söz konusu bilgisayar görmesine geldiğinde durum aynıdır. İnsanın çevresini, 3 boyutlu (3D) olarak görmesini, derinlik hissini algılamasını sağlayan iki gözünün olması dolayısıyla aynı nesnelerin farklı açılardan görüntüsünü alabilmesidir. İki gözden alınan veriler beyin tarafından üst üste bindirilerek derinlik algısı yaratılmasında kullanılır (binoküler). Yapay görme yoluyla derinlik kestirimi de aynı nesnelerin farklı açılardan elde edilmiş resimlerinin işlenmesiyle mümkündür. İnsanlarda bulunan binoküler sistemin incelenmesi ile stereo görüş sistemleri modellenmiş, görüntü işleme teknikleri ile derinlik algılaması işlemi

*İlgili Yazar: Hüseyin Kutlu, Adres: Besni Meslek Yüksekokulu, Bilgisayar Kullanımı Bölümü, Adıyaman Üniversitesi, Adıyaman Türkiye. E-mail: hkutlu@adiyaman.edu.tr, Phone: +904163110422 Fax: +904163110424

gerçekleştirilmiştir [1]. Stereo görüş uygulamaları hareketli nesne algılama, robot yöngüdümü, teleoperasyon ve sanal gerçeklik gibi bir çok alanda kullanılmaktadır [2].

Bu çalışmanın ilk bölümünde günümüzde yaygın olarak kullanılan derinlik algılama hakkında bilgi verilmiştir. İkinci bölümünde ise bu tekniklerden Stereovision ve benzer bir mantıkla çalışan Kinect (prime sence) ile derinlik algılama uygulamalarının nasıl gerçekleştirileceği üzerinde durulmuştur. İki çalışmada da Tsukuba stereo görüntüleri kullanılarak elde edilen sonuçlar gösterilmiştir.

2. Derinlik Algılama Yöntemleri

Derinlik algılama, 3 boyutlu görme (3D), Stereo Vision temelde aynı olayı ifade etmektedirler. Bilinen iki boyutlu resim veya video tek bir kamera ile elde edilir. Tek kamera ile ulaşılan resme veya videoya iki boyutlu görüntü (monocular görüntü) denir. Gerçek dünyadaki görüntüler üç boyutlu olup binoküler görüntü olarak isimlendirilir. Görüntü işleme teknikleri ile resmin veya videodaki nesnelerin üçüncü boyutunu veya derinliğini bulma tekniklerine ise derinlik algılama denir.

Tüm derinlik algılama teknikleri temelde stereo vision (ikili görme) mantığına dayansada aşağıda günümüzde yaygın olarak kullanılan derinlik algılama yöntemlerine değinilmiştir.

2.1. Lazer Üçgenleştirme

Bu modelde, genellikle bir çizgi lazer kaynağı ve bir kamera kullanılır. Sistemin çalışabilmesi için, 3 boyutlu olarak algılanmak istenen nesne kontrollü bir şekilde hareket ettirilmelidir. Kamera sürekli fotoğraf çekerek lazer çizgisi ışığının yansıma değişikliğinden faydalanarak, 3 boyutlu nesnenin profilini çıkarır.



Şekil 1. Lazer Üçgenleştirme Tekniği

2.2. Stereo Vision (İkili Görü)

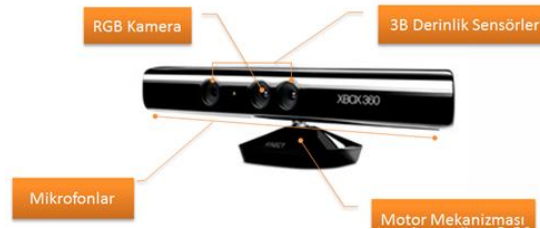
En yaygın olarak kullanılan 3D yöntemidir. İnsan gözü modelinde olduğu gibi, 2 kamera kullanılır. Kameraların arasındaki mesafe ve merkez doğrultu açıları net olarak bilindiğinden, her bir noktanın, her iki kameradaki izdüşümü, basit geometrik hesaplamalar ile bulunabilmektedir.



Şekil 2. Stereo Vision yöntemi

2.3. Kinect

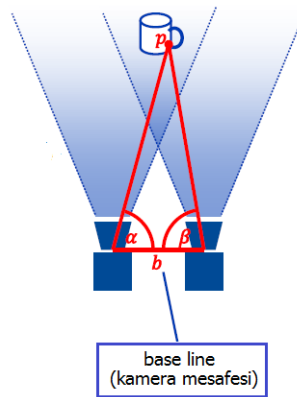
Oyun konsollarının (XBox, PS v.b.) yaygınlaşması ile birlikte genellikle PrimeSense firmasının sensörlerinin kullandığı, Microsoft'un stereovision aracıdır. Kinect cihazının ön tarafında 3 adet göz vardır. Bu üç gözden sağ ve sol taraftakiler 3 boyutlu derinliği algılamakta kullanılan sensorlar iken ortada yer alan göz ise RGB görüntü alabilen bir kameradır. Normal kameralara ilaveten mesafe ölçüm sensörlerinin de bulunduğu, dolayısıyla 3 boyutlu görüntü alınabilen cihazlardır. Mesafe ölçüm sensörleri, genellikle infrared bir kaynaktan, hedef üzerine foton gönderir. Fotonlar hedeften yansıyıp geri gelir. Çok hassas bir zaman ölçümü ile, gelen foton gecikmesinden hedefin bulunduğu konum hassas olarak belirlenir. Genellikle matris şeklinde dizilmiş binlerce sensör kullanıldığından, görüş alanındaki hedef 3 boyutlu olarak hızlı bir şekilde çıkartılmış olur.



Şekil 3. Kinect yöntemi

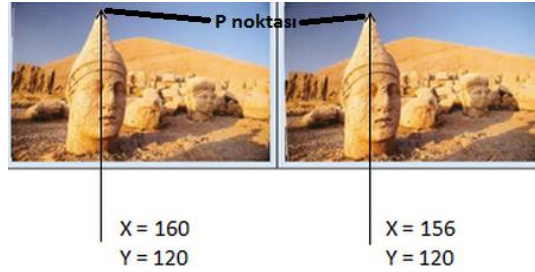
3. Stereo Vision (İkili Görü) Prensibi

Derinlik algılamada temel prensip olan stereo vision aynı nesneye ait, aralarındaki uzaklık bilinen, kalibre edilmiş iki kameradan farklı açılardan alınan resimlerin karşılaştırılması prensibine dayanır.



Şekil 4. Stereo Vision (İkili Görü) Prensibi

Şekil 4. deki resimde kalibre edilmiş (hizaları, yükseklikleri ayarlanmış) iki kameranın aldığı fincan resminin bir pikseli p noktası olsun. α ve β açılarının ve b uzaklığının bilinmesi durumunda geometrik hesaplama ile p noktasının pozisyonu bulunabilir. p noktası bir piksel ise aslında matris olan iki boyutlu bir resimde koordinatları yatayda x. piksel düşeyde y. piksel olmak üzere (x,y) ile gösterilebilir. Farklı açılardan görüntü alan kameralarda p pikselinin x değeri değişim gösterecek başka bir ifade ile kayacaktır. Bu kayma mesafesine Disparity denir. Bu kayma sırasıyla tek gözümüzü kapattığımızda gördüğümüz nesnelere bir miktar kaymasına benzer. Uzaktaki nesnelere için kayma az olacaktır, yakındaki nesnelere için ise kayma daha fazla olacaktır. Bu durumda, stereovision da temel prensip, her bir piksel için Disparity değerini hesaplayabilmektir.



Şekil 5. Sol ve Sağ kameralardan alınan iki resim

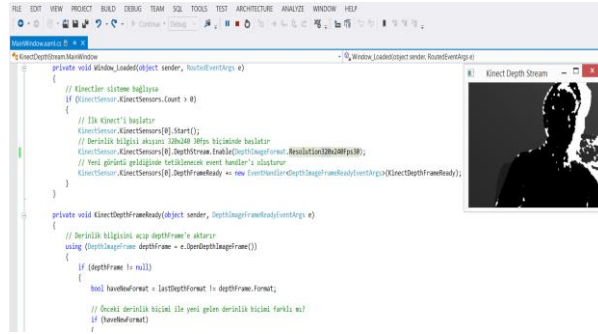
Şekil 5. de p noktası, sağ kameraya daha yakın olduğundan 156. pikselde iken, sol kamerada 160. pikselde. Kameralar kalibre edilmiş olduğundan aynı Y değerlerini vermektedirler. Yine kalibrasyon sonucu, bu 4 piksel kaymanın, ne kadar mesafeye dek düştüğü bulunabilir. Burada, sistemin hızlı çalışması için, sol resimdeki P(160,120) noktasının, sağ resimde nerelerde aranacağını sistem kestirebilmek zorundadır. Kalibrasyon sonucu, sol kameradaki her bir satır (row) sağ kamerada da aynı satıra denk düşecek şekilde resim dönüştürülür (rectification). Son olarak, stereo kamera, hangi mesafeden çalışıyor ise, (10 cm, 30 cm, 1 mt., 2 mt. vs.) yaklaşık olarak, kaç piksel kayma olabileceği bellidir. Böylelikle, sol kameradaki her bir tekil piksel, sağ kamerada tek bir satır üzerinde, dar bir alanda aranır ve hızlıca bulunur.

4. OpenCV ve Kinect ile Derinlik Bilgisine Ulaşma Uygulaması

Çalışma kapsamında yapılan uygulamada açık kaynak görüntü işleme kütüphanesi OpenCV ile farklı açılardan çekilmiş iki resimden derinlik bilgisi çıkarılmış, aynı işlem gerçek zamanlı görüntü üzerinden kinect ile yapılmış ve programın ekran görüntüleri ve algortimaları verilmiştir.

4.1. Kinect ile Gerçek Zamanlı Derinlik Bulma

Çalışma Windows 8.1 işletim sistemine sahip, Intel(R) Core (TM) İ7 - 4510U @ 2.00 GHz 2.6 GHz CPU, 8GB Ram belleğe sahip bir Laptop'a bağlı XBOX 360 KINECT cihazı ile yapılmıştır. Uygulamada kinect kameranın özellikleri kullanılarak görüntü çerçeve çerçeve alınır ve her çerçevenin piksel piksel derinlik değerleri bulunur.



```

private void Window_Loaded(object sender, RoutedEventArgs e)
{
    // Kinectler sisteme başlatılır
    // KinectSensör.KinectSensör.Count > 0)
    {
        // İlk Kinect'i başlatılır
        KinectSensor.KinectSensör[0].Start();
        // Derinlik bilgisi akışı 30Hz'de başlatılır
        KinectSensor.KinectSensör[0].DepthStream.EnableDepthFormat(Resolution30HzFps30);
        // Yeni görüntü geldiğinde her iki Kinect sensörüne de başlatılır
        KinectSensor.KinectSensör[0].DepthFrameReady += new EventHandler<DepthFrameReadyEventArgs>(KinectDepthFrameReady);
    }

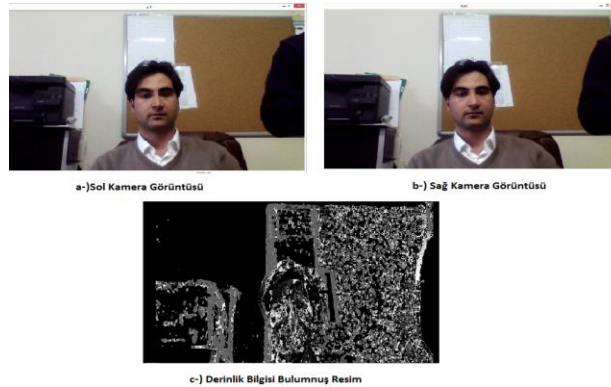
    private void KinectDepthFrameReady(object sender, DepthFrameReadyEventArgs e)
    {
        // Derinlik bilgisini açıp depthFrame'e aktarılır
        using (DepthImageFrame depthFrame = e.OpenDepthImageFrame())
        {
            if (depthFrame != null)
            {
                // Yeni derinlik bilgisi = LastDepthFormat != depthFrame.Format;
                // Eğerki derinlik bilgisi ile yeni gelen derinlik bilgisi farklı mı?
                if (NewDepthFormat)
            }
        }
    }
}

```

Şekil 6. Kinect ile derinlik bilgisine ulaşma

4.2. OpenCV ile Derinlik Bulma

Çalışma Windows 8.1 işletim sistemine sahip, Intel(R) Core (TM) İ7 - 4510U @ 2.00 GHz 2.6 GHz CPU, 8GB Ram belleğe sahip bir Laptop'a bağlı iki adet kameraya bağlı bilgisayarda açık kaynak kodlu görüntü işleme kütüphanesi OpenCV 2.4.10 sürümü Microsoft Visual Studio 2012'ye entegre edilerek C++ programlama dilinde gerçekleştirilmiştir.



Şekil 7. OpenCV Stereo Vision ile derinlik bilgisine ulaşma

Uygulamada OpenCV'nin State komutuyla eşikleme ve blok yapısı işlemleri gerçekleştirilmiş, StereoBM(Sol resim, Sağ Resim, Disparity) komutu ve parametreleri ile çıkış resmi elde edilmiştir.

5. Sonuç

Yapılan çalışmada stereo vision görüntü almada kinectin OpenCV'ye göre kamera kalibrasyonu ve komut kütüphanesinin kullanımı açısından daha avantajlı olduğu görülmüştür. OpenCV'de iki adet kamera kullanımının bellek kullanımını açısından dezavantaj olduğu görülmüştür.

Kaynaklar

- [1] Short N.J., "3-D Point Cloud Generation from Rigid and Flexible Stereo Vision Systems", Yüksek Lisans Tezi, Virginia Polytechnic Institute, Computer Engineering, 2009.
- [2] Lipnickas A. Knys A. "A Stereovision System for 3-D Perception", Electronics and Electrical Engineering, 2009.
- [3] Mustafa Sarı (2014, 14 Temmuz). 3D Görüntü İşleme Teknikleri, Erişim Tarihi:15.02.2015, <http://www.mavis.com.tr/blog/?p=2201>
- [4] Bhatti, A., "Stereo Vision", Intech, 2008.
- [5] KB Özütemiz, A Hacinecipoglu, AB Koku. " Self-learning road detection with stereo vision", SIU, 2013
- [6] MA Akkuş, " Analysis Of Border Ownership Cues And Improvement Of Depth Prediction Using Border Ownership", Yüksek Lisans Tezi, Middle East Technical University, 2014
- [7] M. S. Landy, L. T. Maloney, E. B. Johnston, M. Young, "MEASUREMENT AND MODELING OF DEPTH CUE COMBINATION: IN DEFENSE OF WEAK FUSION", Vision Research, Volume 35, Issue 3, February 1995, Pages 389–412