

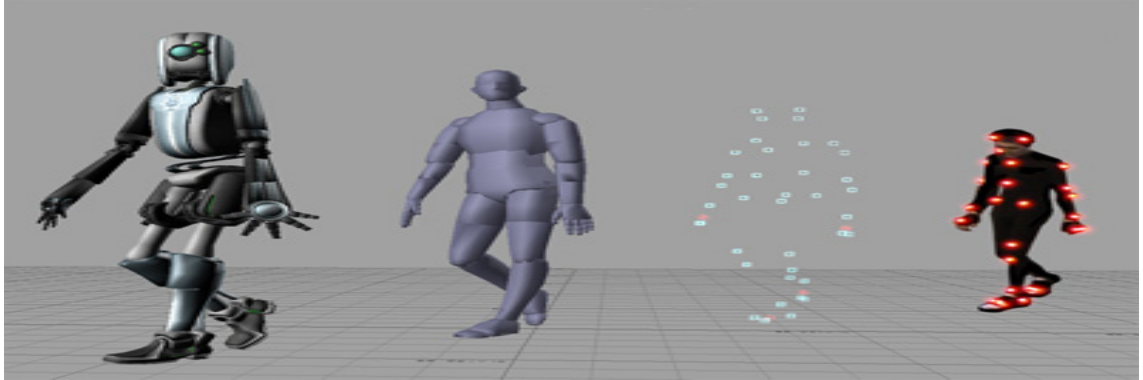


T.C Kütahya Sağlık Bilimleri Üniversitesi

Motion Capture 3D: Gerçek Zamanlı Hareket Yakalama

Ders Sorumlusu: Emre Güngör
İlknur Koparır 2118121010

10.03.2024



1 Giriş

Hareket yakalama teknolojisi (motion capture), yüzyılı aşkın bir süre içerisinde insanlığın geliştirdiği teknolojinin etkileri ile büyük ölçüde gelişim göstermiştir. Günümüzde sanat, eğlence ve teknoloji dünyasında önemli bir yere sahip olan motion capture, insanların gerçek dünyadaki hareketleri dijital ortama aktarma ihtiyacından doğmuştur. Özellikle sinema, video oyunları, reklam endüstrisi gibi alanlarda, gerçekçi hareketlerin oluşturulması ve karakterlerin doğal bir şekilde davranması büyük bir talep görmesinden dolayı bu teknolojilerin kullanım alanları gelişmektedir. Hareket halindeki bir obje veya bir canlının üç boyutlu koordinat ve açısal değişimlerinin çeşitli yöntemlerle bir alıcıya iletilmesi olarak tanımlanan hareket yakalama teknolojisi [1], işaretleyici tabanlı ve işaretleyici tabanlı olmayan sistemler olarak ikiye ayrılmıştır. İşaret tabanlı hareket yakalama sistemleri manyetik veya optik vericilerin insan vücudunun önemli eklem bölgelerine yerleştirilerek hareket bilgisinin üretilmesidir. İşaret tabanlı olmayan hareket tanıma sistemlerde ise verici doğrudan figür olarak karşımıza çıkmaktadır. Kameralar, insan formunu arka plandan ayırt eden algoritmalar ve derin öğrenme teknikleri sayesinde verilerden hareket bilgisi çıkarılmaktadır. Bu teknolojiler sayesinde daha hızlı ve hassas hareket yakalama sistemleri geliştirilerek, gerçek zamanlı etkileşimli uygulamaların performansı artırılmakta ve kullanıcı deneyimi iyileştirilmektedir.



(a) Manyetik hareket yakalama sistemi.



(b) İşaretleyici tabanlı olmayan hareket yakalama sistemi.

2 Literatür Taraması

Motion capture teknolojisinin kökenleri, eski çağlardan başlayarak, insanların hareketlerini taklit etme ve canlandırma ihtiyacıyla başlamıştır.

1906'da Amerikalı mucit Thomas Edison'un yardımı ile Amerikan gazetesi karikatüristi James Stuart Blacton'un, Şekil 2'de görüldüğü üzere ilk çizgi filmi olan Humorous Phases Faces'i (Komik Yüzlerin Mizahi Evreleri) gösterime girmiştir. Çizgi filmde, kara tahta üzerinde, bir şapka ile oynayan palyaço ve çemberden atlayan bir köpek de dahil olmak üzere, elle çizilmiş karakterler bulunmaktaydı. Hareket ettirip durdurma tekniği kullanılan filmde, çizimler kendi başlarına tamamlanıyor ve sonra hareket eden çizimler gibi görünmekteydi. Bu teknik izleyiciler tarafından komik bulunmuş ve ilgi ile karşılanmıştır [2]



Şekil 2: Humorous Phases Faces

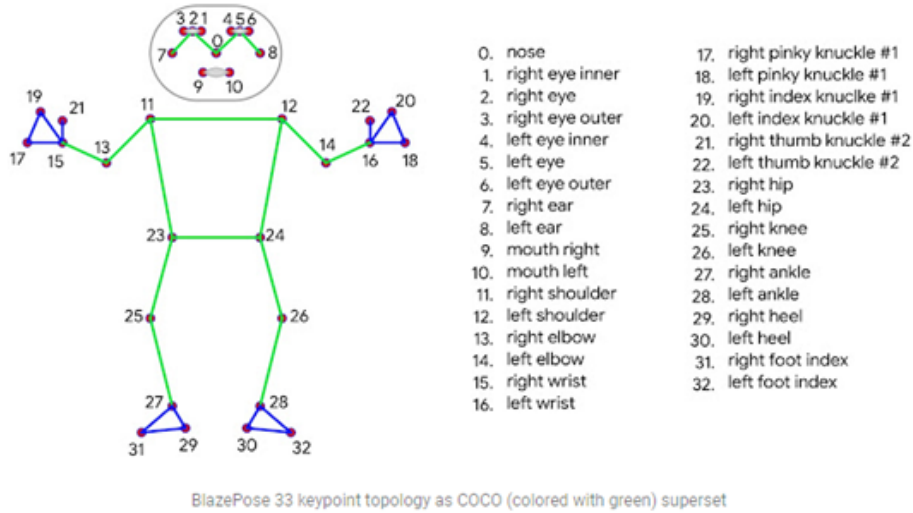
1959'da Harrison, potansiyometrelerle (ayarlanabilir dirençler) donatılmış bir elbiseyi oyuncuya giydirmiş ve oyuncunun tüm hareketlerini bir ekran yardımı ile gerçek zamanlı olarak kaydetmeyi başarmıştır. Bu ilkel bir teçhizatı ancak gerçek zamanlı hareket yakalamanın ilk örneğidir. 1980'lere gelindiğinde animatörler, oyuncuların hareketlerini izlemek için aktif işaretçilerle kaplı vücut giysileri ve avuç büyüklüğünde kameralar kullanmışlardır [2].

Teknolojinin gelişmesiyle birlikte Rosales ve ark. [3], kalibre edilmemiş birden fazla görüş kullanılarak 3B eklemli duruşun (pose) tahmini için bir çerçeve sunar. Bu çerçevede, Uzmanlaştırılmış Haritalama Mimarisi (SMA) adı verilen istatistiksel bir çıkarım yöntemi kullanılır. SMA, kareler ve görünümüler arasında 2B eklem konumları sağlar.

Kakadiaris ve Metaxas [4], dik açılı konfigürasyonlarda çoklu kamera görüntülerini kullanarak insan vücut parçalarının 3B model tabanlı izleme ve şekil tahmini için bir yöntem sunarlar. Daha sonra yaklaşımı, çoklu görüşlü video dizilerinden insan hareketinin tahmini ve buna göre animasyon dizileri oluşturmak için genişlettiler.

3 Metodoloji

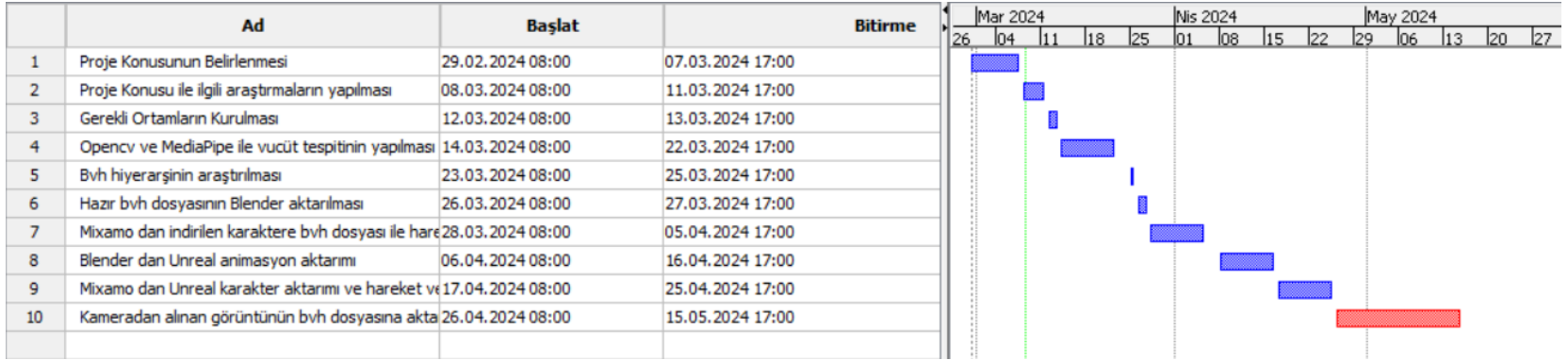
Bu çalışmada, kamera üzerinden alınan insan hareketlerini 3 boyutlu olarak gösterilmesi amaçlanmıştır. Çalışma kapsamında kameradan görüntü almak için OpenCV kütüphanesi kullanılmıştır. Vücut tespiti için ise Google tarafından geliştirilen MediaPipe framework'ü kullanılmıştır. MediaPipe, ücretsiz ve açık kaynaklı bir makine öğrenmesi framework'üdür ve içinde Object Detection, image classification ve test classification gibi birçok model barındırmaktadır. MediaPipe'in Holistic yöntemi, aynı anda yüz, el ve vücut pozisyonlarının takibini mümkün kılmaktadır. Bu yöntem, Şekil 3'de görülebileceği üzere tek bir kare üzerinden 33 noktanın tespit edilmesine izin verirken, geleneksel COCO topolojisinde tek bir kare üzerinde 17 noktanın tespit edilmesini mümkün kılmaktadır. Bu yöntemle, 33 poz, 21 el ve 468 yüz işaret noktası olmak üzere toplamda 543 işaret noktası tespit edilebilir. Mobil cihazlarda bile gerçek zamanlı çalışabilmektedir. Holistic yöntemi çalışırken ilk olarak poz tespiti yapılır, ardından tespit edilen poz işaret noktalarından yola çıkılarak da eller ve yüz bölgesi saptanıp o bölümler ayrıca işaretlenir. En son olarak bulunan tüm noktalar birleştirilir. Böylece vücut tespiti yapılmış olup ilerleyen aşamalarda, MediaPipe ile tespit edilen landmark'lerin bvh dosyasına kaydedilmesi ve bu dosyanın Unreal'da 3 boyutlu bir karaktere uygulanması amaçlanmaktadır. Karakterler öncelikle çevrimiçi ve mocap animasyon veritabanı olan Mixamo'dan fbx formatında indirilip ve daha sonra Unreal ortamına aktarılacaktır.



Şekil 3: MediaPipe'da 33 vücut parçasının konumu

3.1 Sistem Tasarımı

Şekil 4: GANTT CHART



4 Beklenen Sonular

Hareket yakalama teknolojisi son yıllarda biyomekanik, yaya navigasyonu, eğitim ve simülasyon, sanal gerçeklik ve karakter animasyonu alanlarında oldukça önem kazanmıştır. Bu alışmada da kameradan alınan görüntüdeki hareketlerin tespit edilip bu hareketlerin 3 boyutlu olarak unreal ortamına aktarılması amaçlanmaktadır.

Kaynaa

- [1] İ. H. ÖZKİRİŞÇİ, “Motion capture teknolojisinin hareketli afişlerde kullanımına örnek bir alışma,” *Fırat Üniversitesi Sosyal Bilimler Dergisi*, vol. 32, no. 3, pp. 1027–1041, 2022.
- [2] Ö. Ü. S. ERDEM, “Sanali geree dönüştürmede hareket yakalama teknolojisi,” 2021.
- [3] R. Rosales, M. Siddiqui, J. Alon, and S. Sclaroff, “Estimating 3d body pose using uncalibrated cameras, 2001, pp.”
- [4] Y. Desmarais, D. Mottet, P. Slangen, and P. Montesinos, “A review of 3d human pose estimation algorithms for markerless motion capture,” *Computer Vision and Image Understanding*, vol. 212, p. 103275, 2021.