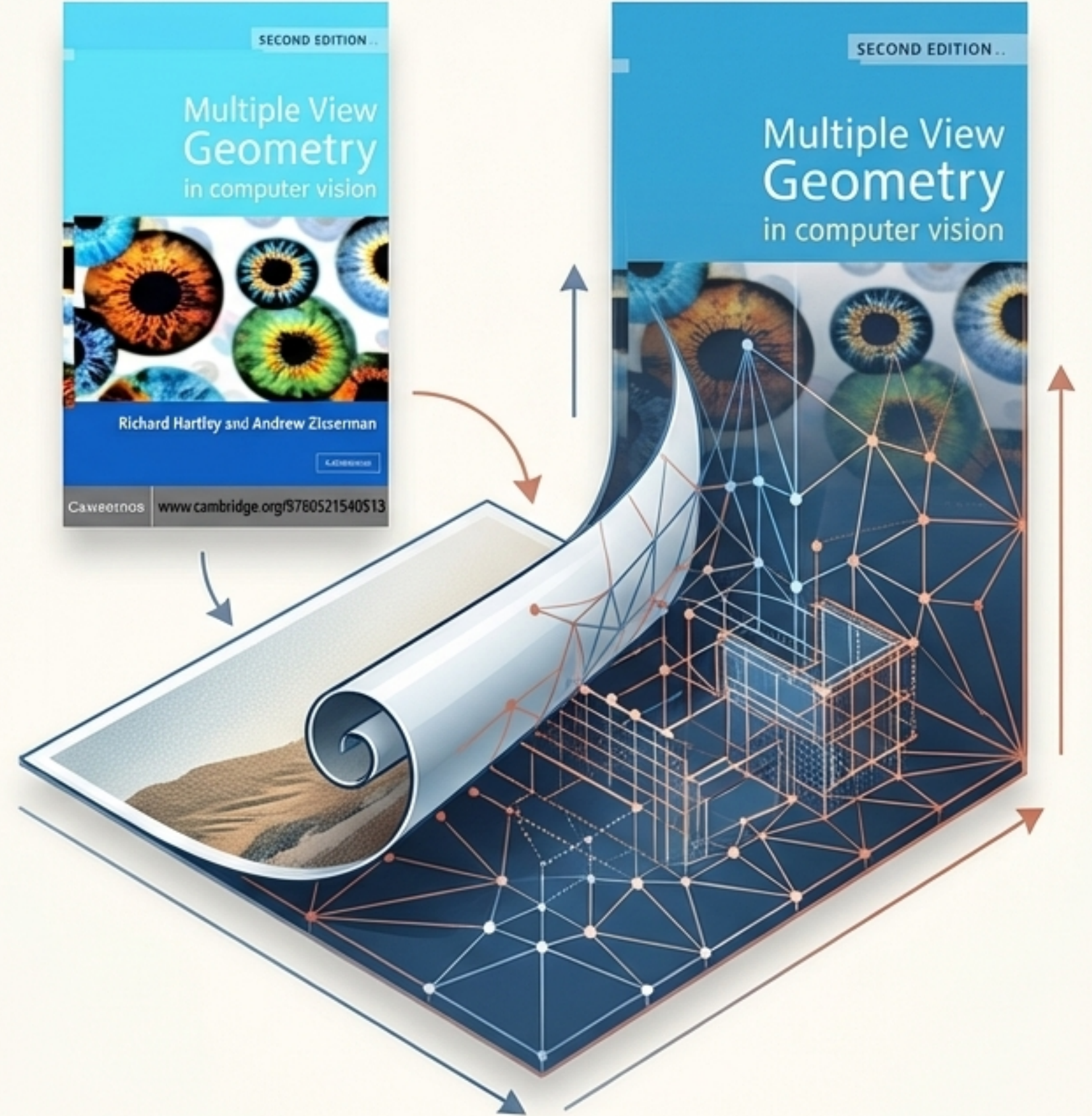


# Boyutları Aşmak: 2D Piksellerden 3D Dünyalara

## Bilgisayarla Görme ve Çoklu Görünüm Geometrisi (Multiple View Geometry)

İnsan beyni dünyayı üç boyutlu algılar. Peki bir bilgisayar, derinlik bilgisini kaybetmiş iki boyutlu düz bir fotoğrafa bakarak gerçek dünyanın 3D yapısını nasıl yeniden inşa edebilir? Cevap, piksellerin izini süren projektif geometride saklıdır.



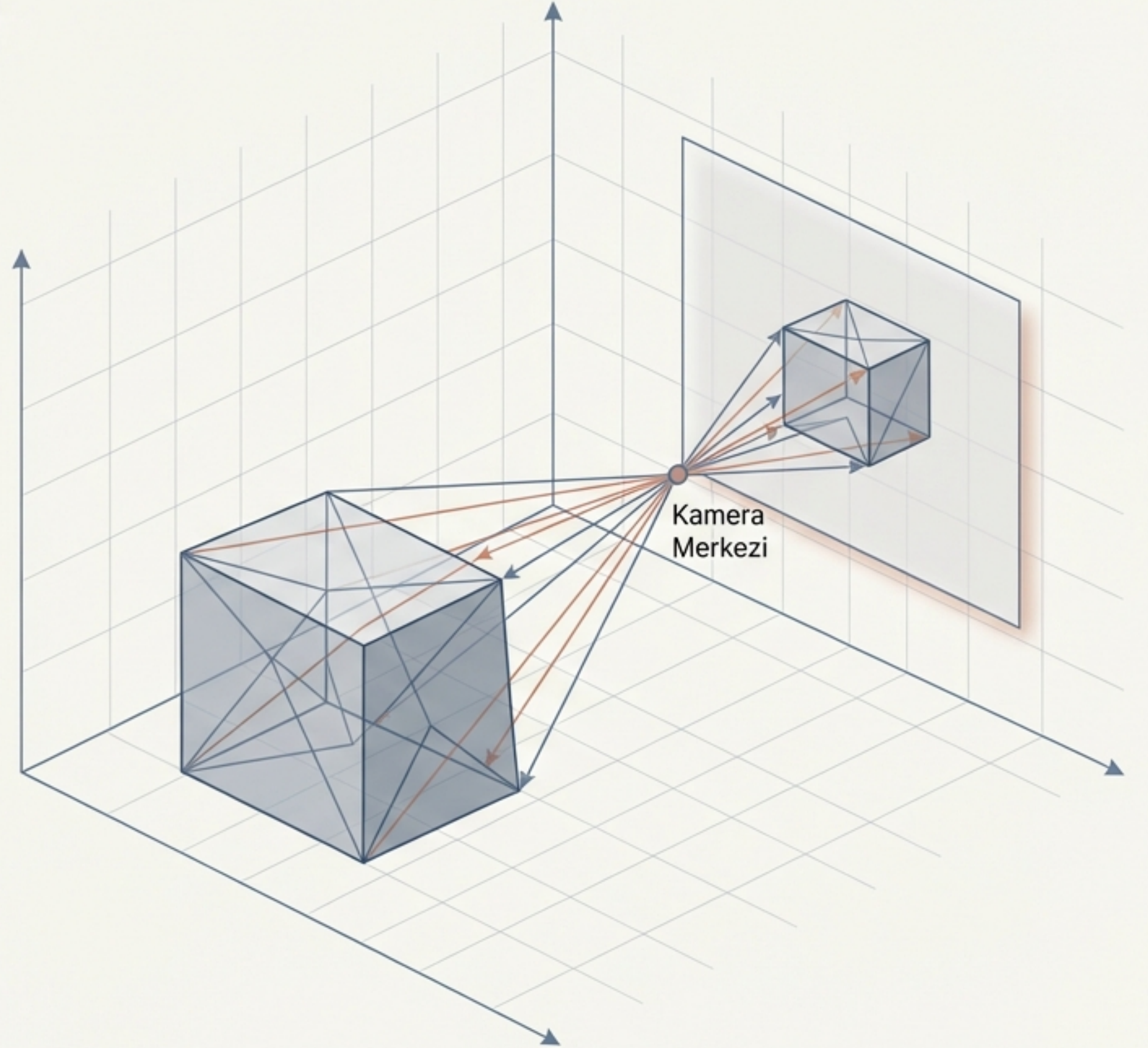


# Problem: Boyut Düşüşü ve Merkezi İzdüşüm

**Kayıp Boyut:** 3D dünyadan 2D görüntüye geçerken derinlik (Z eksen) kaybolur.

**Kamera Merkezi:** Bir görüntüdeki tüm pikseller, uzaydaki tek bir noktadan (kamera merkezi) geçen ışınların (ışık hüzmelerinin) sonucudur.

Gerçek dünya (3D) ile fotoğraf (2D) arasındaki ilişki, "merkezi izdüşüm" (central projection) ile modellenir. Görüntüdeki bir piksel, uzayda kameraya doğru gelen sonsuz sayıdaki noktanın bir özetidir.





# Neden Projektif Geometriye İhtiyacımız Var?

Lisede öğrendiğimiz Öklid geometrisinde (açılar, uzunluklar) paralel çizgiler asla kesişmez. Ancak bir fotoğrafa baktığımızda paralel çizgilerin ufukta birleştiğini görürüz.

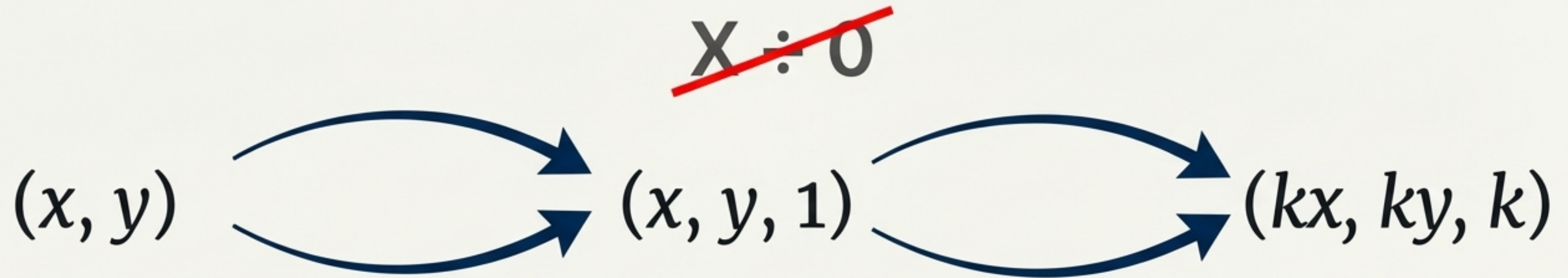
Bilgisayarların bu görsel gerçeği anlaması için “sonsuzdaki noktaları” (ideal points) kucaklayan yeni bir matematiğe, **Projektif Geometriye** ihtiyacı vardır.

Projektif dönüşümlerde şekiller (daire elipse dönüşebilir) ve açılar bozular, ancak **doğrusallık (straightness)** daima korunur.





# Sonsuzluğu Hesaplanabilir Kılmak: Homojen Koordinatlar



Bilgisayarlar sonsuzlukla başa çıkamaz. Çözüm, koordinat sistemine fazladan bir boyut eklemektir.  $(x, y)$  yerine  $(x, y, w)$  kullanmak,  $w=0$  olduğunda bilgisayarın bilgisayarın çökmesini engeller ve “sonsuzluktaki ufuk noktalarını” sıradan, hesaplanabilir rakamlara dönüştürür.

## **\*\*Sihirli Matematik:**

$w$  (veya  $k$ ) değeri sıfır olduğunda, uzaydaki “sonsuzluk” temsil edilebilir hale gelir.



# Geometrilerin Hiyerarşisi ve Projektif Bozulma



Kameraların iç yapısını bilmeden yapılan bir 3D modelleme "Projektif Rekonstrüksiyon" ile sonuçlanır. Sistem doğru görünse de, uzay rastgele esnemiş ve bozulmuştur. Gerçek dünyaya (Öklid geometrisine) ulaşmak için hiyerarşik bir yolculuk gerekir.

**1. Projektif:** Sadece doğrular korunur. (En bozuk kupa)



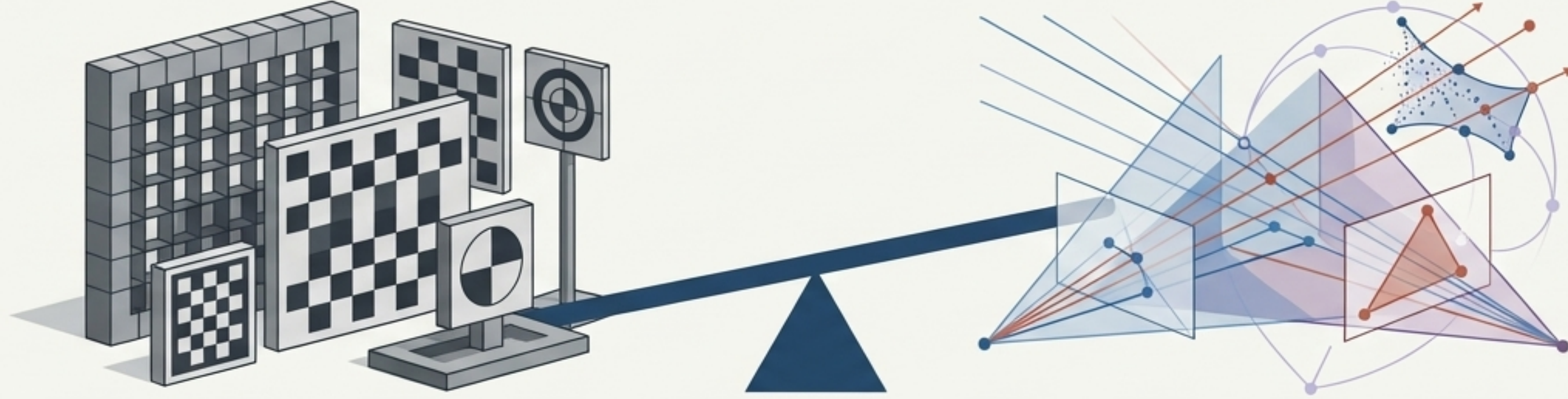
**2. Afin:** Paralellik keşfedilir ve korunur.



**3. Öklid:** Açılar, uzunluklar ve gerçek şekil ortaya çıkar. (Ortadaki kusursuz kupa)



# Paradigma Değişimi: Kalibrasyona İhtiyaç Yok



Kalibrasyon Objeleri

Görüntü İçi Geometri

Eskiden 3D modelleme için her kameranın odak uzaklığı gibi iç parametrelerinin (11 parametre) önceden kalibre edilmesi şarttı. Hartley ve Zisserman'ın öncülük ettiği devrim ise şuydu: Sadece fotoğraflardaki noktaların eşleşmesini kullanarak, kameranın özelliklerini **görüntünün kendisinden** hesaplamak!

*Uncalibrated Framework:* Kameraları değil, doğrudan görüntülerin içindeki geometriyi kullan.

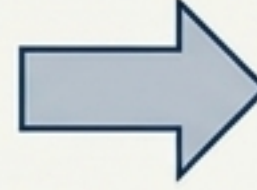


# Tekil Görünüm Geometrisi (1-View)

Tek bir fotoğraftan 3D dünya yaratabilir miyiz?

Evet, eğer sahnede paralel çizgiler (örneğin fayanslar, dolap kenarları) varsa.

Ufuk çizgileri ve kaybolan noktalar (vanishing points) analiz edilerek, sadece bir tablodan bile sahnenin Afin ve Öklid yapısı çıkarılabilir.



Rönesans ressamlarının kusursuz perspektifi, yüzlerce yıl sonra bilgisayarlar tarafından 3D modellere dönüştürülebilir.



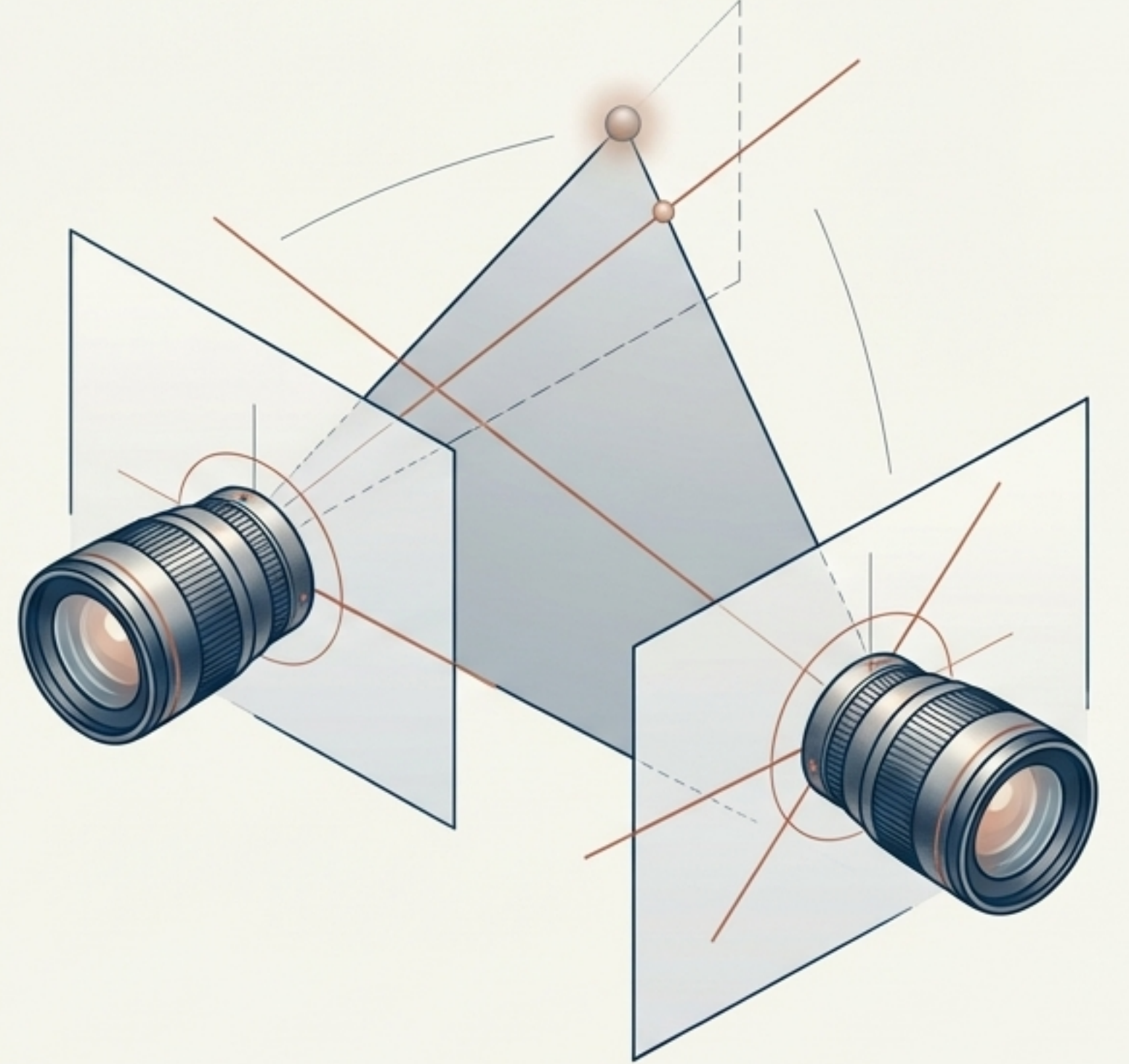
# İki Görünüm Geometrisi: Sihir Başlıyor

$$\text{Formül: } \mathbf{x}'^T \mathbf{F} \mathbf{x} = 0$$

İki farklı açıdan çekilen görüntüler, **Temel Matris (Fundamental Matrix - F)** adı verilen 3x3 bir matrisle birbirine bağlanır.

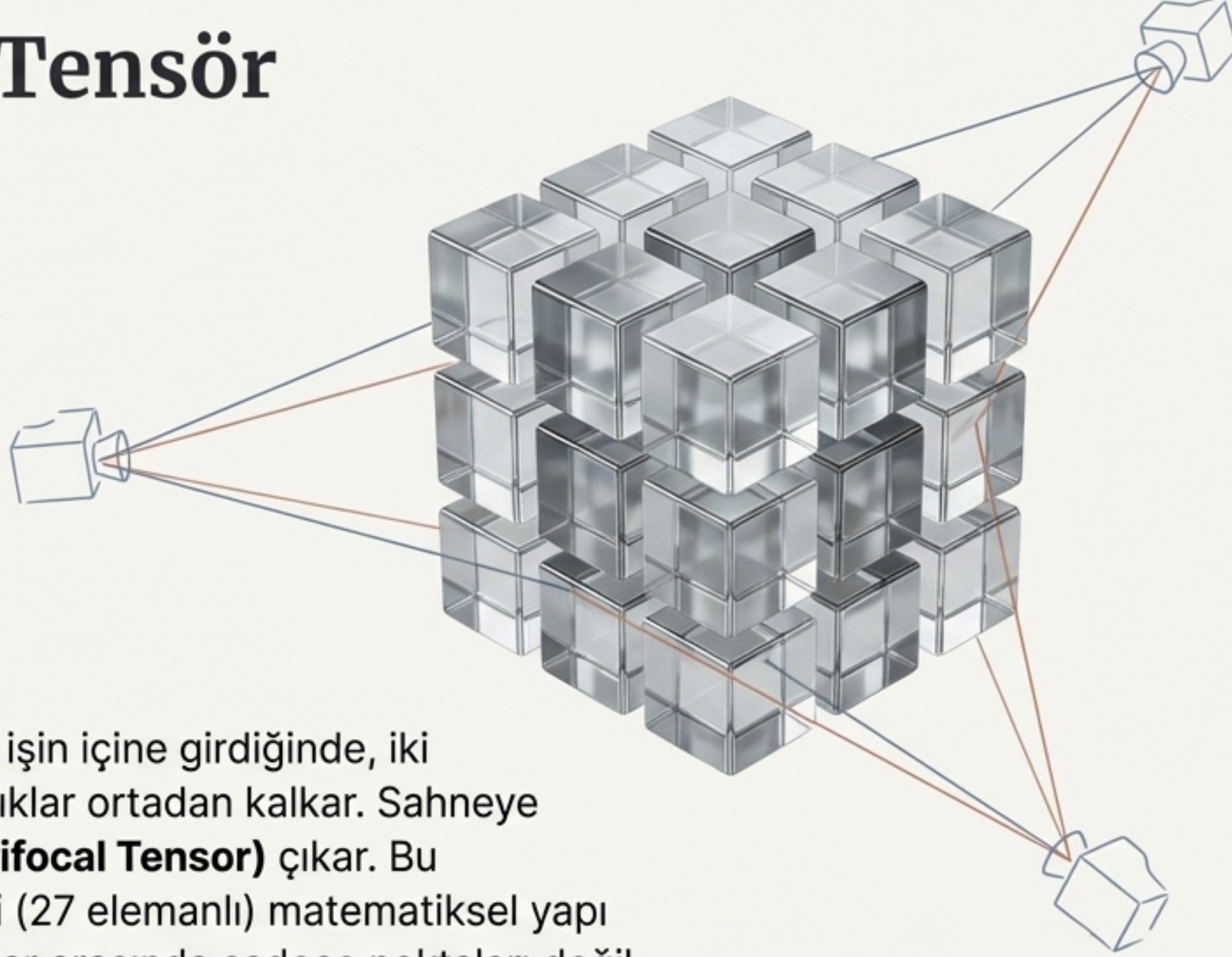
Kameraların kalibrasyonu bilinmese bile, sadece 7 noktanın eşleştirilmesiyle uzaydaki piksellerin 3D konumları "üçgenleme" (triangulation) yapılarak bulunabilir.

F Matrisi, iki kameranın tüm projektif geometrisini içinde barındıran kalbidir.





# Üç Görünüm Geometrisi: Trifokal Tensör



Üçüncü bir kamera işin içine girdiğinde, iki görünümdeki zayıflıklar ortadan kalkar. Sahneye **Trifokal Tensör (Trifocal Tensor)** çıkar. Bu  $3 \times 3 \times 3$  boyutundaki (27 elemanlı) matematiksel yapı sayesinde, kameralar arasında sadece noktaları değil, *çizgileri* de kusursuz bir şekilde eşleştirebiliriz.

Neden 3 Görünüm?  
Doğrusal özellikleri eşleştirebilme gücü ve çok daha yüksek 3D stabilite.

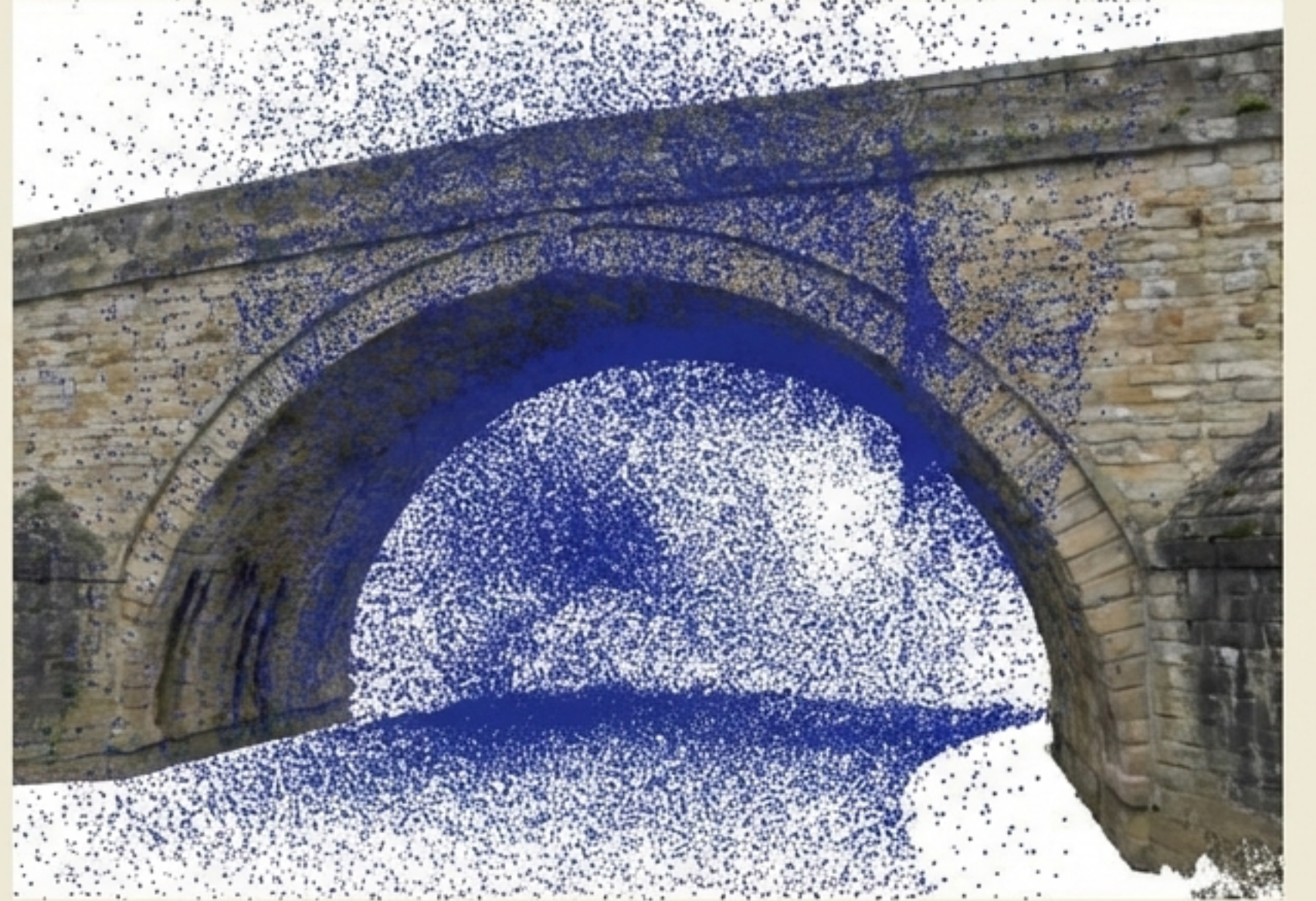


# N-Görünüm ve Demet Dengelemesi (Bundle Adjustment)

Ya elimizde bir videodan alınmış yüzlerce kare varsa?

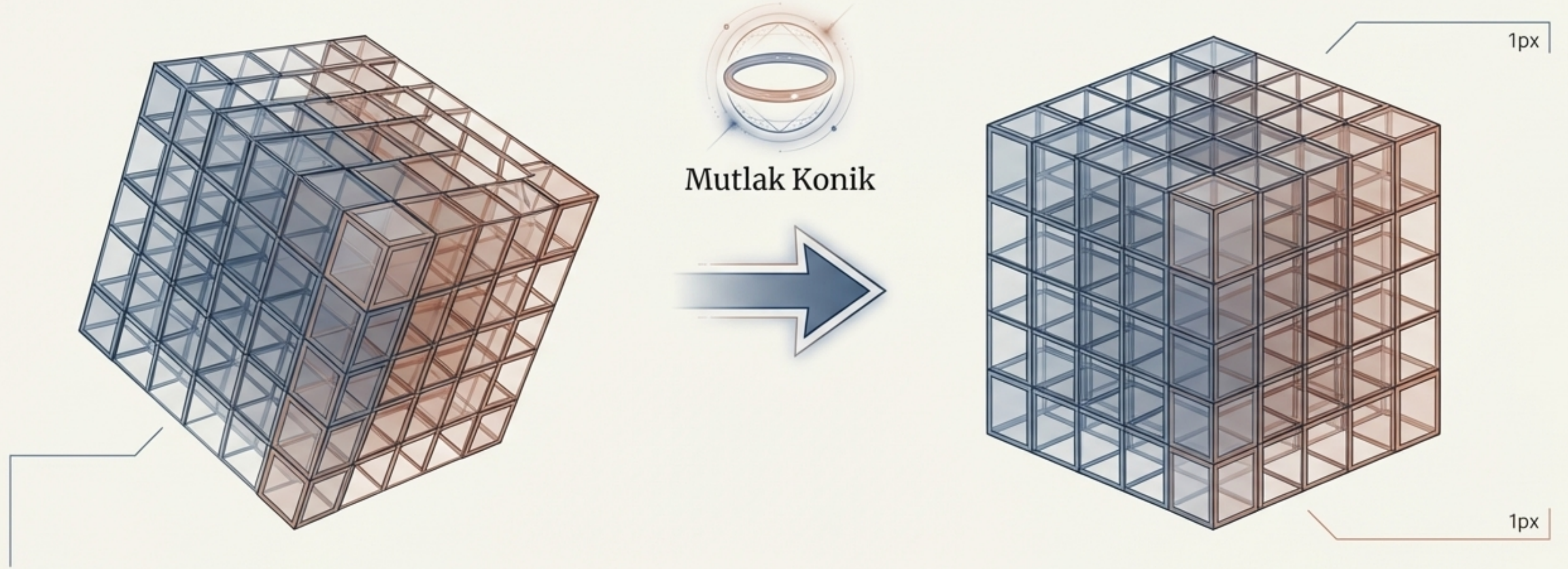
Milyonlarca noktanın ve yüzlerce kamera pozisyonunun aynı anda optimize edilmesi gerekir.

Demet Dengelemesi (Bundle Adjustment) adı verilen iteratif optimizasyon algoritmaları, tüm bu kaosu çözer ve devasa nokta bulutlarını (point clouds) olağanüstü bir hassasiyetle hesaplar.





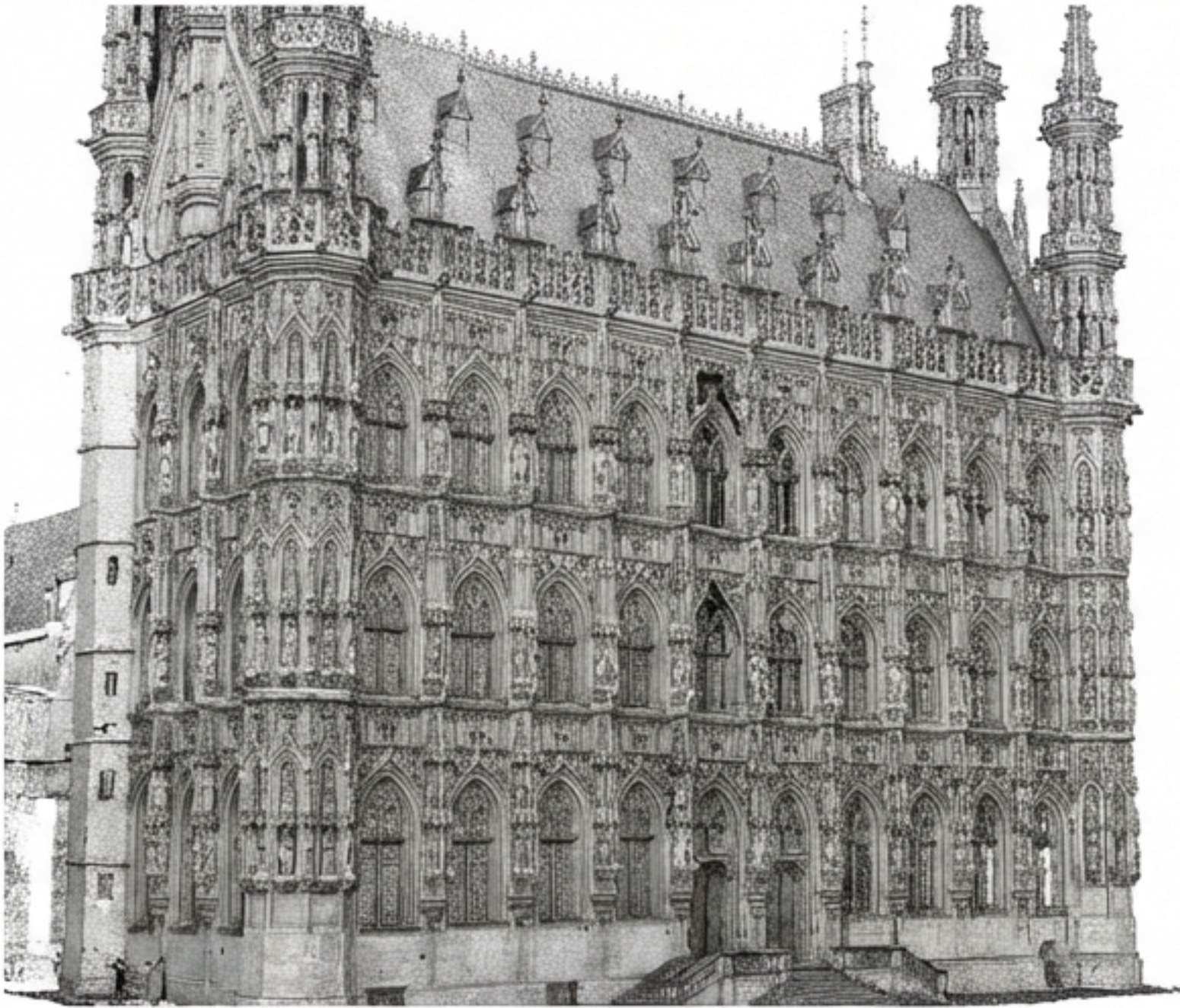
# Otomatik Kalibrasyon (Auto-calibration)



Kamera ayarlarını bilmiyoruz. Sadece piksellerin hareketini biliyoruz. Yine de **gerçek dünya ölçülerine dönebilir miyiz?** Evet. **Piksellerin kare olduğu varsayımı (square-pixel constraint)** gibi ipuçlarını kullanarak, uzayın sonsuzluğunda gizlenen **"Mutlak Konik"** bulunur. Mutlak konik bulunduğunda, kamera kendi kendini kalibre etmiş olur.



# Geometrinin Ödülü I: 3D Grafik Modeller



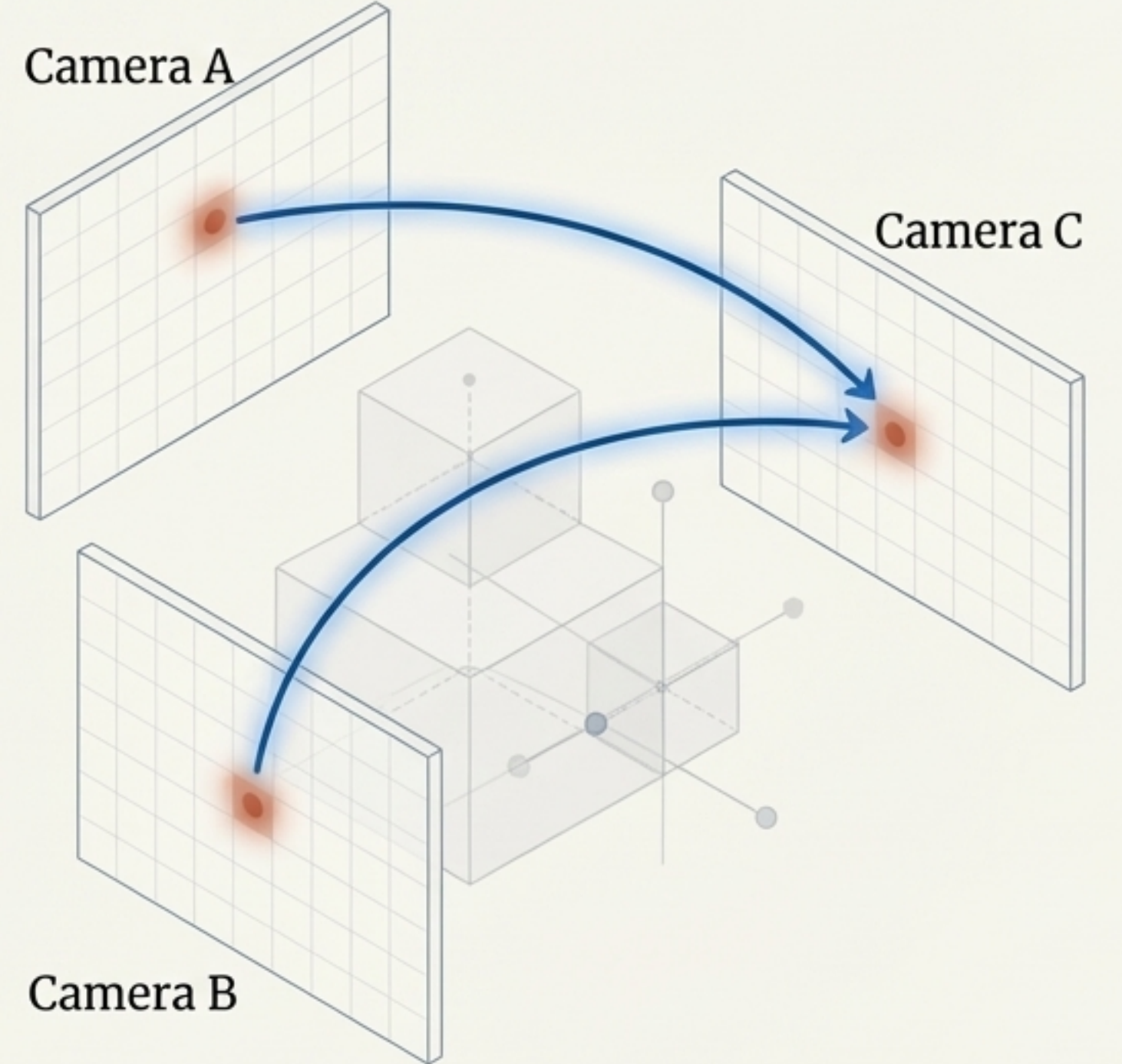
Geometri tamamlandığında geriye görsel bir şölen kalır. Kameraların yerleri hesaplandıktan sonra, eşleşme araması (*correspondence search*) epipolar çizgilerle daraltılır. Sonuç: Sıradan 2D fotoğraflardan otomatik olarak üretilen, yoğun (*dense*), gölgelendirilmiş ve doku kaplanmış fotogerçekçi 3D mimari modellerdir.



# Görünüm Aktarımı (Transfer)

Çoklu görünüm geometrisinin bir diğer süper gücü “Transfer” işlemidir. Bir noktanın iki görüntüdeki yerini biliyorsak, 3D bir uzay modeli inşa etmeye *hiç gerek kalmadan*, çoklu görünüm tensörlerini kullanarak bu noktanın üçüncü bir kamerada tam olarak nereye düşeceğini hesaplayabiliriz.

Hızlı, doğrudan ve sadece pikseller arası matematiksel geçiş.





# Geometrinin Ödülü II: Artırılmış Video (Video Augmentation)

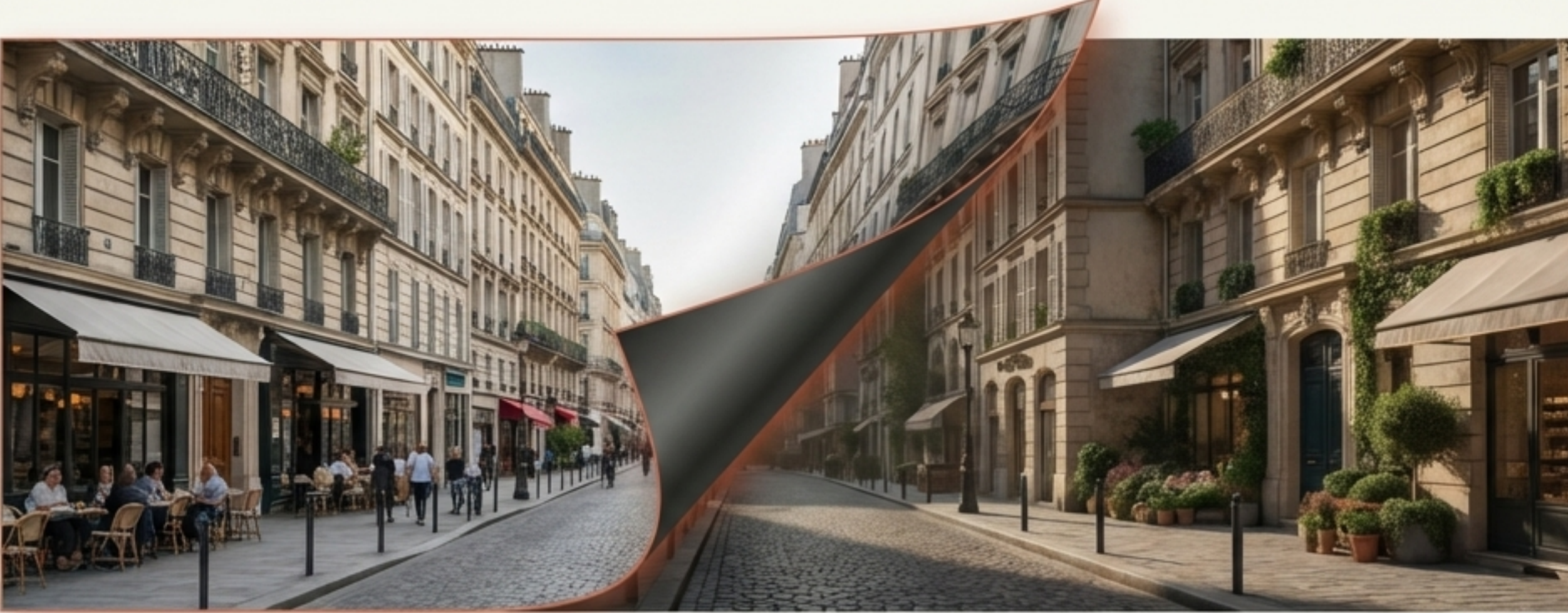
Film endüstrisi ve Artırılmış Gerçeklik (AR), kameranın hareketini gerçek dünyada (Öklid uzayında) kusursuzca bilmek zorundadır. Aksi takdirde eklenen CGI objeler sahnede “kayar”.

Çoklu görünüm geometrisi, el kamerasıyla çekilmiş sıradan bir videonun 3D haritasını çıkararak sentetik karakterlerin (örneğin bir robotun) sahneye fiziksel olarak doğru yerleştirilmesini sağlar.





# Sadece Bakmak Değil, Derinliğı Anlamak



2D piksellerden oluşan bir dünyadan, projektif geometrinin, tensörlerin ve homojen koordinatların rehberliğinde 3D bir evren inşa ettik. Bilgisayarla görme teknolojisi sayesinde bilgisayarlar artık dünyayı sadece iki boyutlu düzlemlerde kaydetmiyor; tıpkı bizim gibi uzayı, derinliğı ve formu **anlıyor**.

Piksellerin ardındaki matematiğı çözdüğümüzde, tek sınırimız hayal gücümüzdür.