Österreichische
Mathematische
Gesellschaft

# MATHE-BRIEF

**April 2020 — Nr. 99**

Herausgegeben von der Österreichischen Mathematischen Gesellschaft
http://www.oemg.ac.at/Mathe–Brief ——— mathe–brief@oemg.ac.at

## „Ultrascharfe“ Fotos?

### Sind Fotografien Zentralprojektionen?

Fotografie übt auf viele Menschen eine Faszination aus – auch auf Mathematiker und insbesondere auf Geometrie-Begeisterte. Im Geometrie-Unterricht sagt man etwas vereinfachend: Fotografien entsprechen Zentralprojektionen (Perspektiven) des Raums. Damit geht eine Reduktion des dreidimensionalen Raums in die zweidimensionale Ebene einher. Als „Beweis“ gilt, dass man aus mehreren Fotografien mit recht guter Genauigkeit die fotografierte räumliche Szene rekonstruieren kann. Voraussetzung ist allerdings, dass qualitativ hochwertige Linsensysteme verwendet werden, die insbesondere geradlinige Kanten exakt geradlinig abbilden.
Auch wenn die Idee für Szenen mit größeren technischen Objekten wie Polyedern (z. B. Zimmer samt Einrichtung oder Gebäude) gar nicht schlecht funktioniert, gibt es doch massive Probleme im Makrobereich, also bei Fotografien von Objekten, die nur wenige Zentimeter groß oder gar kleiner sind.

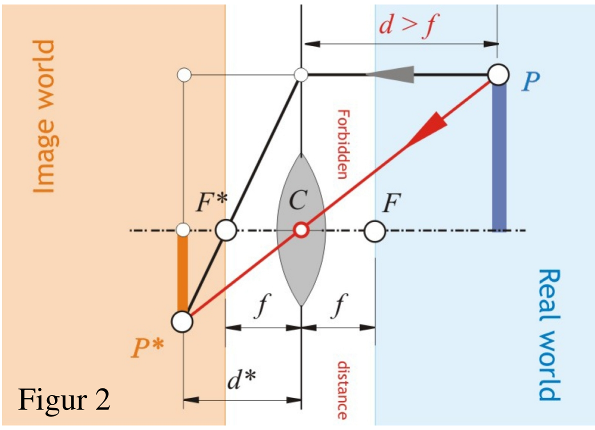
### Ein unmögliches Foto

Figur 1

Die in Fig. 1 abgebildeten knapp 1 cm großen Fliegen unterscheiden sich fototechnisch bei genauerem Hinsehen grundlegend. Das Foto links (mit dem herausgewürgten Verdauungssaft) ist durchgehend scharf. Ein Insektenfotograf wird hier stutzig: Selbst mit den besten Makroobjektiven scheint

so ein Foto nämlich unmöglich – auch bei Verwendung einer teuren Ausrüstung mit speziellen Makroobjektiven, Makroblitzen und der größtmöglichen Blendenzahl, wie beim rechten Foto, wo winzige Wassertröpfchen auf den Komplexaugen zu sehen sind.
Hier treten zwei Fragen auf: Warum kann man ein so kleines Objekt wie eine Fliege nicht durchgehend scharf fotografisch abbilden, und wie scheint es dann doch zu gehen? Dieser Aufsatz soll die Sache mathematisch-geometrisch klären.

**Die Linsengleichung**



Figur 2

In der Physik wird die Wirkungsweise einer Linse (oder eines gut abgestimmten Linsensystems) wie folgt erklärt: Sei $P$ ein Punkt der realen Welt. Er emittiert (reflektiert) nach allen Richtungen Lichtstrahlen. Zwei davon haben beim Durchgang durch das Linsensystem leicht vorhersehbare Eigenschaften. Der *Hauptstrahl* durch das Linsenzentrum $C$ wird nicht gebrochen, während der Strahl parallel zur optischen Achse nach der Brechung durch den Brennpunkt $F^*$ geht. Hinter der Linse (dem Linsensystem) treffen sich die beiden Strahlen – und alle anderen – in einem Bildpunkt $P^*$.

Aus Fig. 2 lässt sich daraus unter Verwendung von ähnlichen Dreiecken die *Linsengleichung* ableiten:

(1) $$\frac{1}{f} = \frac{1}{d} + \frac{1}{d^*}.$$

Dabei ist $f$ die Brennweite, und $d$ bzw. $d^*$ sind die orientierten Abstände des Raumpunkts $P$ bzw. Bildpunkts $P^*$ von der Symmetrieebene durch das Linsenzentrum $C$.
Durch Rotation um die optische Achse kann man so zu jedem Raumpunkt den zugehörigen Bildpunkt ermitteln.

**Die Gaußsche Kollineation**

Setzen wir $d = kf$. Dann ergibt sich aus der Linsengleichung

(2) $$d^* = \frac{k}{k-1} \cdot f = \frac{d}{k-1}.$$

Nach dem Strahlensatz ergibt sich damit für die Abstände eines Raumpunkts $P$ und dessen Bildpunkt $P^*$ vom Zentrum $C$ die einfache Beziehung

(3) $$\overline{CP^*} = \frac{1}{k-1} \cdot \overline{CP}.$$

Auch wenn die geometrische Abbildung $P \mapsto P^*$ in beide Richtungen funktioniert, wird eine Kamera nur jenen Halbraum abbilden können, dessen Punkte weiter als die Brennweite $f$ vor der Linse liegen ($k > 1$): Punkte in der Ebene durch den Punkt $F$ senkrecht zur optischen Achse werden auf Fernpunkte abgebildet, weil dann der Nenner $k-1$ verschwindet.

Es ist leicht zu zeigen, dass die Abbildung $P \mapsto P^*$ *geradentreu* ist: Sei $g$ eine beliebige Gerade des Raums. Diese kann stets als Schnitt zweier spezieller Ebenen $\varepsilon$ und $\varphi$ definiert werden. Dabei ist $\varepsilon$ die Verbindungsebene von $g$ mit dem Zentrum $C$ und $\varphi$ ist jene Ebene durch $g$, die parallel zur optischen Achse ist. $\varepsilon$ geht in sich über ($\varepsilon^* = \varepsilon$), weil man sich die Ebene als Büschel von Hauptstrahlen denken kann. Die andere Ebene $\varphi^*$ kann man sich aus Strahlen parallel zur optischen Achse vorstellen, die in ein Büschel durch den Brennpunkt $F^*$ und die Schnittgerade von $\varphi$ mit der Symmetrieebene übergehen. Das Bild $g^*$ von $g$ ist der Schnitt von $\varepsilon^*$ und $\varphi^*$ und somit eine Gerade. Unsere Abbildung $P \mapsto P^*$ ist folglich eine Kollineation – was ja der Fachausdruck für geradentreue Abbildung ist[1]. Diese Erkenntnis geht auf C. F. GAUSS zurück.

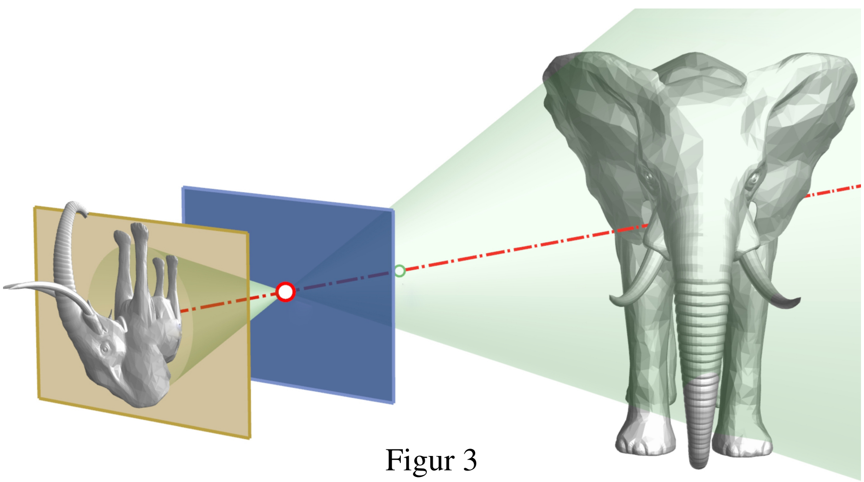
### Der Zusammenhang mit der Fotografie

Mittels der einfachen Formel (3) lassen sich räumliche Objekte, die aus vielen Punkten zusammengesetzt sind, sehr einfach in ebenso *räumliche* Objekte transformieren. Wo bleibt da der Zusammenhang mit der Fotografie, die ja ein *zweidimensionales* Ergebnis liefert?

Betrachten wir wieder einen Punkt $P$ im Abstand $d$ von der Symmetrieebene (Kollineationsebene). Befindet sich die Ebene $\pi$ des Sensors unserer Kamera zufällig im Abstand $d^*$ von der Hauptebene, liegt der Bildpunkt $P^*$ in $\pi$. So gesehen wird also der Schnitt des abzubildenden Objekts mit der „Schärfenebene" durch $P$ im Abstand $d$ parallel zur Kollineationsebene scharf in der Sensorebene abgebildet[2]. Alle anderen Punkte erscheinen mehr oder weniger unscharf.
Das Ausmaß der Unschärfe hängt, wie wir noch sehen werden, von verschiedenen Parametern ab. Ein entscheidender Parameter ist die Größe des Abstands des fotografierten Objekts im Verhältnis zur Brennweite.

### „Elefantenfotografie" vs. „Fliegenfotografie"

Betrachten wir zunächst die Abbildung eines „großen" Objekts aus großer Distanz. Gemeint ist dabei die Relation der Ausmaße des Objekts zur Brennweite $f$ (Fig. 3).



Figur 3

Für Punkte mit großer Distanz $d = k \cdot f$ $(k \gg 1)$ variiert die Bildweite $d/(k-1)$ (Formel 2) nicht allzu sehr.
Das kollineare virtuelle Objekt hinter der Linse wird damit stark abgeflacht sein und dementsprechend wird die Unschärfe jener Punkte, die nicht genau in der Schärfenebene liegen, gering ausfallen. Von einem Elefanten ein bildfüllendes scharfes Foto zu machen, ist also kein Problem.

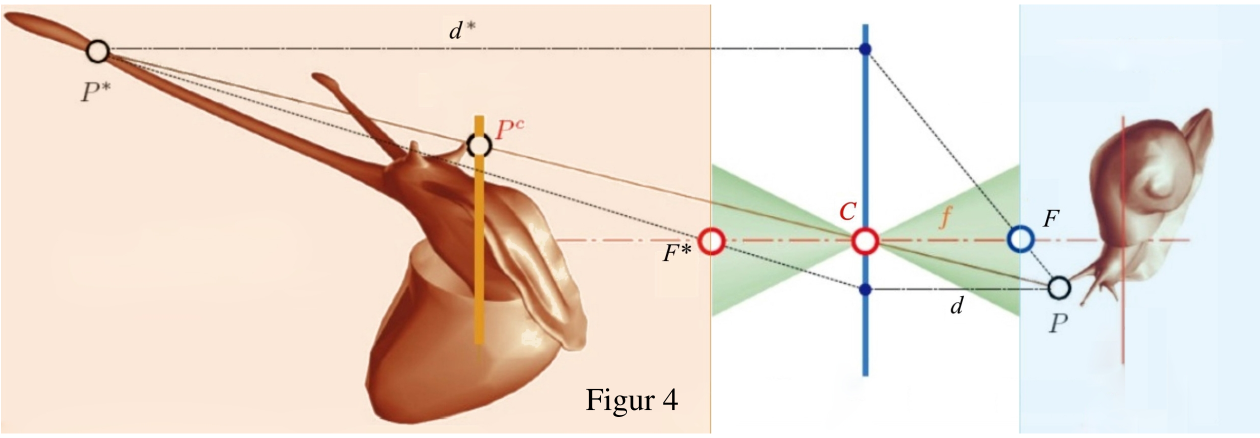
Je kleiner ein zu fotografierendes Objekt ist bzw. je näher dieses Objekt an die „verbotene" Verschwindungsebene rückt (Fig. 4), desto mehr wird der Ausdruck $f/(f-d) = 1/(1-k)$ in der

[1] Für Eingeweihte: Die Abbildung ist eine ganz spezielle *perspektive Kollineation*: Das Zentrum liegt in der *Kollineationsebene* (der Symmetrieebene). So eine Kollineation nennt man *Elation*.

[2] Man hätte mit dem Hausverstand annehmen können, dass in der fotografischen Abbildung all jene Punkte scharf abgebildet werden, die vom Linsenzentrum einen bestimmten konstanten Abstand haben (also auf einer Kugel um $C$ liegen), der vom Abstand $d^*$ der Sensorebene abhängt. Dies ist aber nach der Linsengleichung nicht der Fall, denn besagte Punkte liegen allesamt in einer Ebene, der *Schärfenebene* im Abstand $d$.

Transformationsformel (3) variieren. Dadurch wird die Unschärfe der Punkte zu einem echten Problem.

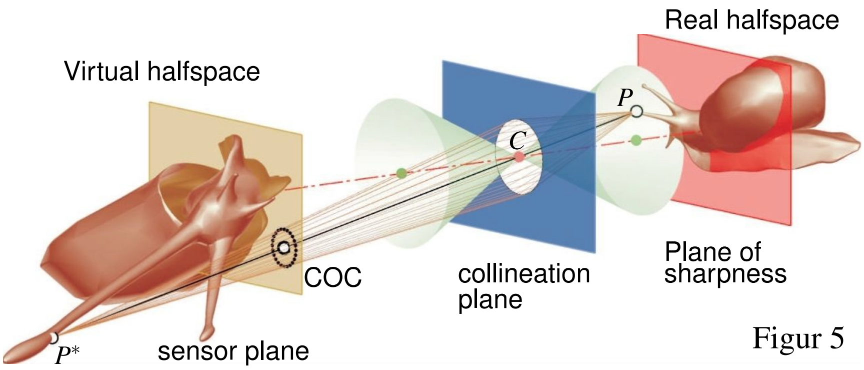
Anmerkung: Kleine Brennweiten $f$ wirken sich offensichtlich positiv auf die Schärfentiefe aus, weil dann eine Fliege oder Schnecke im Verhältnis zu $f$ größer wird. Kameras mit kleinen Sensoren haben eine entsprechend kleinere Brennweite[3].



Figur 4

**Geometrie vs. Physik**

In der Geometrie scheint die Sache einfach: Man schneide den Lichtstrahl durch das Linsenzentrum mit der Sensorebene. Physikalisch gesehen ist das natürlich nicht so: Ein einziger Lichtstrahl reicht nicht aus, um den Sensor zu belichten. Man wird daher in der Hauptebene eine kreisförmige Öffnung – die Blende – einbauen müssen. Alle Lichtstrahlen, die von einem Raumpunkt $P$ ausgehen, liegen dann innerhalb eines schiefen Kreiskegels durch die Öffnung, der in einen ebenfalls schiefen Kreiskegel gebrochen wird (Fig. 5).



Figur 5

Die Gesamtheit aller Lichtstrahlen in diesem gebrochenen schiefen Kreiskegel belichtet die Sensorebene nur dann punktförmig, wenn $P$ in der *Schärfenebene* liegt. Bei allen anderen Punkten ergibt sich am Sensor ein sog. Unschärfekreis (englisch: COC = *circle of confusion*).

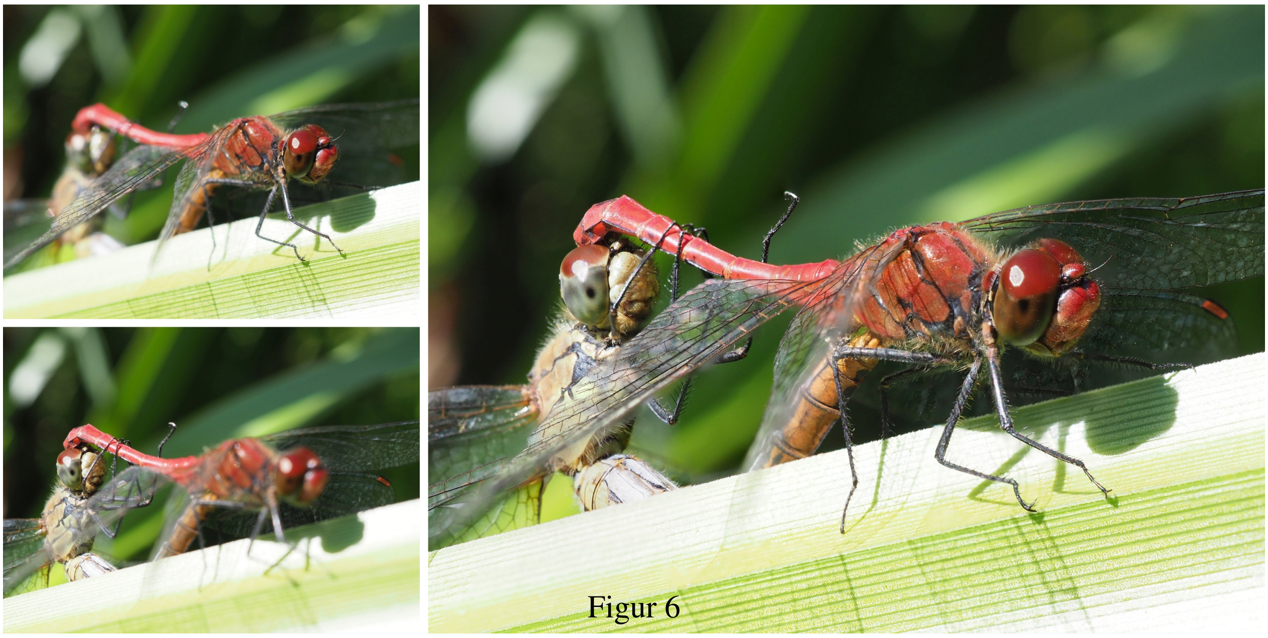
Man könnte nun annehmen, dass man nur genügend Licht verwenden müsse (Blitz), um die Blendenöffnung möglichst klein halten zu können (man spricht dann von einer hohen Blendenzahl). Dies ist aber nur bis zu einem gewissen Limit möglich (die Blendenöffnung sollte jedenfalls mehr als 1 mm betragen). Verkleinert man weiter, bekommt man es mit den Welleneigenschaften des Lichts

[3]Oft wird bei Werksangaben die Brennweite eines Objektivs z. B. mit 100 mm KB-äquivalent angegeben, das heißt, das Objektiv hätte bei einer Kleinbild-Sensorgröße von 24 mm × 36 mm eine Brennweite von 100 mm. Wenn der Sensor aber z. B. nur die Ausmaße 6 mm × 9 mm beträgt, erreicht man mit einer Brennweite von nur 25 mm denselben Bildeindruck. Im speziellen Fall liegt ein *Crop-Faktor* 4 vor. Mittlerweile kann man sogar mit guten Smartphones wegen deren extrem kurzen Brennweiten und entsprechend winzigen Sensoren (mit noch viel höheren Crop-Faktoren) erstaunlich scharfe Makrofotos machen. Das Problem ist hier allerdings die hohe Anzahl der Pixel auf kleinstem Raum, die unweigerlich Qualitätseinbußen nach sich zieht.

zu tun: Es kommt zu einer Beugung am Rand der Öffnung, was eine unangenehme *Beugungsunschärfe* erzeugt. Das Optimum liegt dann bei einer von den Objektivherstellern angegebenen *förderlichen Blende*. Fotografen wissen: Das Überschreiten der förderlichen Blendenzahl mindert die Bildqualität.

**Focus stacking**

Bei der Elefantenfotografie halten sich die Probleme in Grenzen und wenn man als Fotograf einen Punkt anvisiert, der in etwa am Ende des ersten Drittels der gewünschten Entfernungsbandbreite liegt, wird das Bild zufriedenstellend scharf sein[4].
In der Makrofotografie aber ist das Schärfenproblem gravierend – insbesondere, wenn es sich nicht nur um künstlerische, sondern um wissenschaftliche Aufnahmen handelt.



Figur 6

Mittlerweile hat sich eine Technik namens *focus stacking* etabliert, die im Wesentlichen wie folgt funktioniert: Die Kamera nimmt – in möglichst kurzem Zeitabstand – mehrere Bilder der Szene auf, wobei der Abstand der Schärfenebene variiert wird. Auf diese Weise bekommt man eine Bilderserie, in denen der Reihe nach verschiedene Schichten des Objekts scharfgestellt sind. Die Theorie der Bildbearbeitung ist mittlerweile schon sehr fortgeschritten und entsprechende Software kann scharfe von unscharfen Bildpunkten unterscheiden.

Aus der gesamten Bilderserie wird dann in einem letzten Schritt *ein* scharfes Bild erzeugt. Wie gut das sogar ohne Stativ funktionieren kann, soll Fig. 6 zeigen: Links sind exemplarisch das erste und das letzte Bild einer solchen Serie zu abgebildet, rechts ist das Endergebnis zu sehen[5].

[4]Oft hat man als künstlerischer Fotograf das gegenteilige Problem: Man will ja gezielt mit Unschärfen arbeiten. Hier empfiehlt sich die Verwendung größerer Brennweiten und der sogenannten *Offenblende*.

[5]Ein zusätzlicher Vorteil der Methode ist, dass man im Normalfall kein „unendlich“ scharfes Bild erhält, sondern dass eben eine gewisse Zone des Raums scharfgestellt ist. Unscharfe Hintergründe ermöglichen ein Freistellen des abzubildenden Objekts und verhindern, dass der Blick des Betrachters durch unnötige Details abgelenkt ist.

## Die Schärfenebene tastet das Objekt ab

Für jede Position der Sensorebene (Abstand $d^*$) gibt es in der Gaußschen Kollineation somit genau eine Schärfenebene (Abstand $d$), deren Position sich aus der Linsengleichung (1) ergibt:

(4) $$d = fd^*/(d^* - f).$$

Bei einer marktüblichen Kamera bleibt die Position der Sensorebene fest und die Position des Linsenzentrums $C$ wandert beim Scharfstellen auf der optischen Achse vor und zurück. Werden sehr weit entfernte Punkte scharfgestellt, befindet sich $C$ im Abstand $f$ vor der Sensorebene (mit $d^* = f$ ist $d = \infty$). Fotografieren wir ein zweidimensionales Gebilde – also z. B. eine Zeichnung – das sich in einer Ebene parallel zur Sensorebene im Abstand $s = d + d^*$ befindet und stellen scharf, dann nimmt das Zentrum $C$ jene Position im Abstand $d^*$von der Sensorebene ein, die es zu berechnen gilt.

Mit Formel (4) gilt

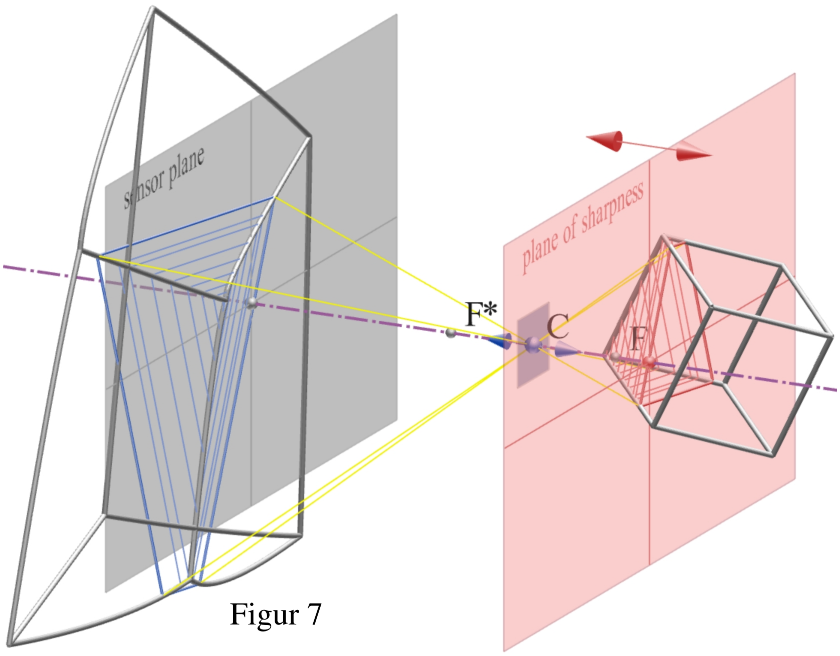
(5) $$s = d^{*2}/(d^* - f) \quad \text{bzw.} \quad d^{*2} - sd^* + sf = 0$$

Die zweideutige Lösung dieser quadratischen Gleichung ist

$$d^* = \frac{s}{2} \pm \sqrt{\frac{s^2}{4} - sf}.$$

Damit der Ausdruck unter der Wurzel nicht negativ ist, muss $s \geq 4f$ sein. Dies ist wegen der Voraussetzung $d > f$ tatsächlich immer der Fall (bei $d = d^* = 2f$ fallen die Lösungen zusammen). In der Tat sind beide Lösungen stets zulässig, wobei man bei der Berechnung einer Serie von Kamerapositionen sinnvollerweise immer bei *einem* Vorzeichen bleiben wird.

Liegt nun das Linsenzentrum im berechneten Abstand $d^*$ vor der festen Sensorebene $\pi$, bildet sich unser zweidimensionales Gebilde erstens durchgehend scharf am Sensor ab, und zweitens erscheint es dort *ähnlich*, also perspektivisch unverzerrt, wenn auch nicht mehr in wahrer Größe.




Figur 7

Kehren wir nun wieder zu dreidimensionalen Objekten zurück. Wenn wir von so einem Objekt eine Serie von Fotos bei gezielter Veränderung der Distanz des Linsenzentrums machen, erhalten wir Fotos, auf denen jeweils eine Schichtenlinie des Objekts scharf *und im Wesentlichen unverzerrt* – nämlich nur skaliert – abgebildet wird. Um die Sache zu illustrieren, wurde in Fig. 7 in einer Simulation des Vorgangs ein Würfel auf diese Weise in Schichten, die Dreiecke, Vierecke, Fünfecke oder Sechsecke sein können, abgebildet. Wenn wir die erhaltenen Bilder nun einfach nur übereinander legen, kommen wir allerdings auf seltsam veränderte Perspektiven (siehe dazu auch Fig. 8 links). Geraden – etwa die Kanten des Würfels – werden dabei gekrümmt abgebildet (man kann zeigen, dass es Parabeln sind). Das liegt natürlich daran, dass durch die unterschiedliche Distanz des Linsenzentrums von der Sensorebene die Schichtenlinien proportional zu dieser Distanz skaliert sind.

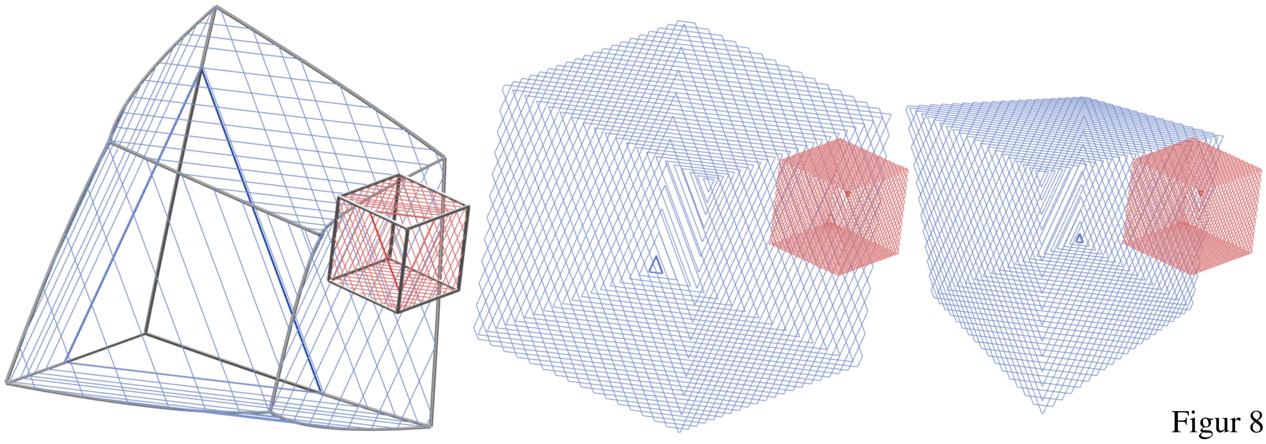
**Umrechnung des Scan-Vorgangs in Normal- und Zentralprojektion**

Genau genommen haben wir auf diese Weise unser Objekt dreidimensional gescannt – wenn auch unter der Einschränkung, dass nur die vom jeweiligen Zentrum sichtbaren Schichtenlinien erfasst wurden. Wir wollen nun noch gekonnt skalieren, um die Verhältnisse im während des Scan-Vorgangs sichtbaren dreidimensionalen realen Raum rechnerisch unter Kontrolle zu bekommen.

Sei $t$ eine Strecke in der Schärfenebene, also im Raum, und $t^*$ die zu ihr parallele Strecke in der Sensorebene $\pi$, dann gilt nach dem Strahlensatz $t : d = t^* : d^*$. Mit dem Skalierungsfaktor

$$\lambda = d/d^* \tag{6}$$

kann man also die wahre Länge $t$ aus der Bildlänge berechnen. Skaliert man alle Bilder der Serie mit dem von Bild zu Bild wechselnden Faktor, erhält man dann eine *Normalprojektion* des Objekts (vgl. dazu Fig. 8 Mitte). Das ist insofern besonders, als man mit einem *einzelnen Foto niemals eine Normalprojektion* erreichen kann, es sei denn durch ein astronomisches Teleskop mit fast unendlicher Brennweite[6].



Figur 8

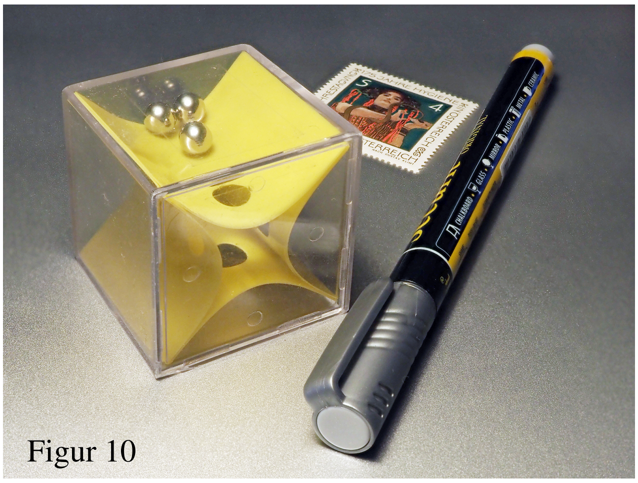
Um eine exakte Zentralprojektion (Perspektive) zu erreichen, wo ausschließlich eindeutig erfasste Punkte verwertet werden, wird man daher als Zentrum jene Position wählen, an der das Linsenzentrum war, als die letzten noch sichtbaren Punkte des abzubildenden Objekts scharf erschienen ($d = d_m$ maximal $\Rightarrow d^* = d_m^*$ minimal). Das $i$-te Bild wird dann nicht nur gemäß Formel (6) mit dem Faktor $\lambda_i = d_i/d_i^*$, sondern zusätzlich mit dem Faktor $d_m^*/d_m$ skaliert:

$$\mu = \frac{d_i}{d_i^*} \cdot \frac{d_m^*}{d_m} \tag{7}$$

Wenn man jetzt die Bilderserie der Software zur Erkennung der scharfen Pixel übergibt, erhält man mit graphischer Genauigkeit ein scharfes und perspektivisch korrektes Bild, wie man in Fig. 8 rechts sehen kann.

[6]Eine Einschränkung ist allerdings gegeben: Selbst aus mehreren Positionen aus der optischen Achse kann man nicht immer so viel auf einer Oberfläche eines Objekts erkennen wie bei einer echten Normalprojektion. Man denke etwa an ein Raumschiff, das sich dem Mond nähert. Von ihm aus wird man niemals 50% der Mondoberfläche sehen können, wie von der Erde aus.

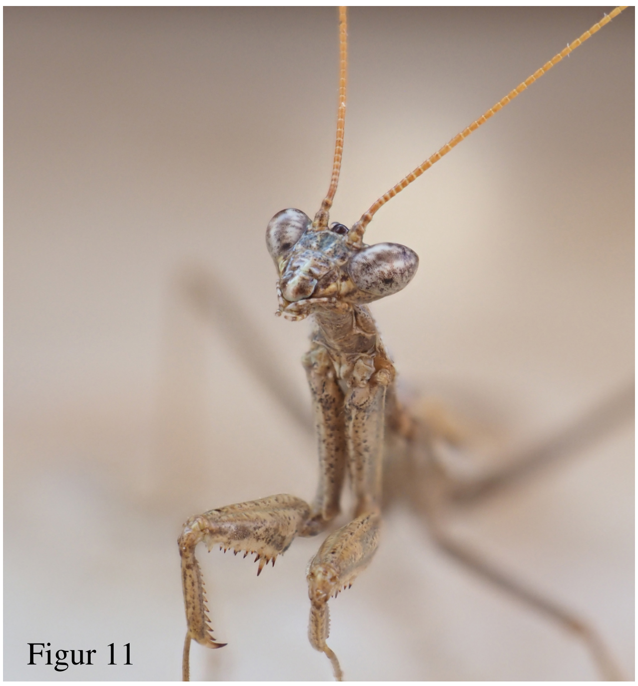
Figur 9


Figur 10

So besteht das Foto der quadratischen Briefmarke in Fig. 9 relativ gut die strengen geometrischen Tests, die man anwenden kann, um die Korrektheit einer Perspektive zu testen, ebenso das „geometrische Stillleben“ in Fig. 10. Das ist deshalb wichtig, weil man in der Praxis ja nicht geometrische Miniaturfiguren, sondern Lebewesen oder andere Gegenstände der Natur fotografiert. Die Briefmarke *bildfüllend* scharf zu fotografieren, ist eine klassische Aufgabe der Makrofotografie, während das Stillleben im Verhältnis dazu schon eine mittelgroße Szene und entsprechend leichter durchgehend scharf zu fotografieren ist.

**Ausblick**

Vom mathematisch/geometrischen Standpunkt ist noch anzumerken, dass man mit dem vielfach fotografierten Objekt auf Grund der räumlichen Erfassung noch einiges mehr „anfangen“ kann, etwa 3D-Modelle erstellen. Jedenfalls sollte man mit focus stacking im Makrobereich Ergebnisse erreichen können, die unter Laborbedingungen an jene von Laserscannern herankommen – wobei die Technologie einfacher, schneller und billiger ist.


Figur 11

Wenn die Objekte Kleintiere sind, kommt das Problem dazu, dass diese eher selten bewegungslos verharren, zumindest aber Fühler oder einzelne Glieder bewegen (Fig. 11). Hier sollte die Bildserie in längstens einer Zehntelsekunde abgearbeitet sein, was wohl bei der rasanten technologischen Entwicklung in wenigen Jahren möglich sein wird. Derzeit brauchen handelsübliche Kameras noch etwa eine halbe Sekunde für eine volle Serie von 8 bis 10 Bildern, wobei der Flaschenhals beim Zeitverbrauch nicht das Abspeichern der Bilder, sondern das Neu-Justieren des Fokus ist.

Vom ästhetischen Standpunkt aus reichen für bemerkenswerte Insektenfotos aber oft auch zwei, drei Bilder des Tiers. In Fig. 11 war wichtig, die Augen und die Greifzangen der Gottesanbeterin scharf zu bekommen, beim restlichen Körper stört es nur marginal, dass dieser unscharf erscheint.

Georg Glaeser