Bildunterschriften zu Objekterkennung

- **Folie 3**: Trotz der Variation von Farben und Formen (äußerer Kreis) sowie verschiedenen Blickwinkeln (mittlerer Kreis) ist es für einen Menschen relativ einfach möglich, Bilder von Motorrädern anhand einiger Identifikationsmerkmale (innerer Kreis) als solche zu erkennen.

- **Folie 7**: Die ROC illustriert Kombinationen von True Positive Rate (Sensitivität) und False Positive Rate (eins minus Spezifizität), wodurch die Leistungsfähigkeit verschiedener Objekterkennungsalgorithmen verglichen werden kann. Raten (rote strichlierte Linie) ist dabei als Referenz durch die Punkte auf der ersten Mediane (z.B. B) beschrieben – alle darüber liegenden Algorithmen (z.B. A) klassifizieren besser, alle darunter liegenden (z.B. C) schlechter. Ein perfekter Objekterkennungsalgorithmus liegt am Ende der TPR-Achse, da bei ihm keine Falscherkennungen auftreten.

- **Folie 8**: Bei der ROC-Kurve wird die True Positive Rate über der False Positive Rate aufgetragen, wodurch die Leistungsfähigkeit eines Objekterkennungsalgorithmus bei variabler Falscherkennungsrate beschrieben werden kann. Der Kurvenverlauf pro Algorithmus ist streng monoton, da eine erhöhte Falscherkennungsrate in jedem praktisch relevanten Fall zur Verbesserung der Erkennungsrate führt. Reines Raten (schwarz strichliert) ist die Referenz, über der ROC-Kurven von Algorithmen mindestens liegen sollten. Knapp darüber liegende Kurven (z.B. grün) werden als schwach eingestuft, wohingegen nahezu sprungförmige an den oberen Koordinatensystemgrenzen (z.B. blau) als fast ideal eingestuft werden.

- **Folie 9**: Beim Objekterkennungsalgorithmus nach Viola und Jones werden objektcharakteristische Hell-Dunkel-Unterschiede im Originalbild (links unten) durch entsprechende Merkmale (oben) modelliert, die in den dafür vorgesehenen Gesichtsregionen (Mitte unten bzw. rechts unten) gut übereinstimmen.

- **Folie 10**: Die Merkmaltypen, die im Objekterkennungsalgorithmus von Viola und Jones verwendet werden, sind kontrastreiche rechteckförmige Hell-Dunkel-Muster.

- **Folie 11**: Jedes Bildmerkmal (Beispiel links) wird per Schiebefensterverfahren über das Bild bewegt (Mitte), um an den entsprechenden Übereinstimmungsstellen (Beispiel rechts) einen Treffer auszulösen.

- **Folie 13**: Die Fläche des Rechtecks R mit dem linken oberen Eckpunkt $\begin{pmatrix} x_l \\ y_t \end{pmatrix}$ und dem rechten unteren Eckpunkt $\begin{pmatrix} x_r \\ y_b \end{pmatrix}$ in einem Bild (angedeutet durch das äußere Rechteck) kann mit Hilfe eines Integralbildes einfach berechnet werden. Für den Sonderfall eines $1 \cdot 1$ Pixel großen Rechtecks kann durch diese Berechnung der dazugehörige Pixelwert im Originalbild rekonstruiert werden.

- **Folie 15**: Bei einer Kaskadierung von Detektoren muss jeder Detektor eine positive Rückmeldung (1) geben, damit der Bildausschnitt als Gesicht erkannt wird. Gibt ein Detektor eine negative Rückmeldung (0), werden die weiteren Detektoren in der Kaskade übersprungen.

- **Folie 17**: Das Schiebefensterverfahren wird auf das Originalbild (links) sowie auf herunterskalierente Versionen desselben (Mitte bzw. rechts) angewandt.

- **Folie 18**: Vom Originalbild (links) werden in verschiedenen Auflösungen jeweils Integralbilder (Mitte) berechnet, auf die Detektorkaskaden (rechts) angewandt werden können.

- **Folie 19**: Die Pixel eines Originalbildes (grau) sollen nach der Herunterskalierung weniger dicht platziert sein. Dadurch beinhaltet ein Pixel des herunterskalierten Bildes (weiß) mehrere Pixel des Originalbildes – das Einzugsgebiet jedes Pixels mit Fläche $4a^2$ ist mit einer dicken schwarzen Linie markiert. Das linke Bild zeigt einen ganzzahligen Skalierungsfaktor in beiden Richtungen (die Position der weißen Pixel relativ zu den grauen ist konstant), das rechte Bild einen nicht ganzzahligen (die Position der weißen Pixel relativ zu den grauen ändert sich mit der Pixelposition im Bild).

- **Folie 21**: Die bilineare Interpolation des Pixels $P = \begin{pmatrix} x \\ y \end{pmatrix}$ (grün) aus den Original(-nachbar-)pixeln $Q_{ij} = \begin{pmatrix} x_i \\ y_j \end{pmatrix}$, $i, j \in \{1, 2\}$ (rot) erfolgt durch (links) zweimalige lineare Interpolation in x-Richtung (blau, $R_j = \begin{pmatrix} x \\ y_j \end{pmatrix}$, $j \in \{1, 2\}$ aus $\overline{Q_{1j}Q_{2j}}$) und anschließende lineare Interpolation in y-Richtung ($P = \begin{pmatrix} x \\ y \end{pmatrix}$ aus $\overline{R_1R_2}$). Alternativ (rechts) kann das zu interpolierende Pixel (schwarz) aus der gewichteten Summe der zwischen dem Pixel und dem jeweils gegenüberliegenden Nachbarpixel eingeschlossenen Flächen berechnet werden.

- **Folie 22**: Je nach Pixelabstand (x-Achsen-Wert) vom zu interpolierenden Pixel werden die Original(-nachbar-)pixel mit dem Funktionswert der roten Kurve an dieser Stelle gewichtet. Verschiedene Werte von b (links $b = 2$, rechts $b = 3$) führen zu unterschiedlichen Gewichtungen sowie einer veränderten Anzahl berücksichtigter Nachbarpixel (vier bei Lanczos-2, sechs bei Lanczos-3).

- **Folie 24**: In der von Lienhart und Maydt vorgeschlagenen Erweiterung des Objekterkennungsalgorithmus nach Viola und Jones werden neben horizontalen und vertikalen Bildmerkmalen auch diagonale (1. (c) und (d), 2. (e)-(h) und 3. (b)) verwendet.