时，在多模态大模型机器人应用方向上创业的公司需要找准细分领域，明确用户需求，解决好特定问题，避免在“通用性”上过度纠结，明确”**盈利**“而非”研究“为目标，使用基础大模型针对特定机器人任务进行优化即可。使用自然语言作为人机交互媒介，要避免机器人**反馈过多无效信息**，注重提升**交互体验**。
Google作为在多模态大模型-机器人领域里走在最前面的公司，其内部是否会再次孵化子公司，或有相关人员重新创业，值得关注。

除此之外，并无更多将多模态大模型应用于机器人领域的公司。总体上来说这一领域仍处于**技术储备阶段**，需要观望，但非常值得期待。

机器人领域应用外，还有一些大模型公司也值得关注，此处不再展开介绍：Boson.ai，Mu Li与其导师Alex Smola从Amazon离职，投身大模型创业，CLIPr，Twelve Labs，Lightricks，Jasper，Stability.ai，Hugging Face；小冰科技在2022年11月进行了A+轮11亿人民币的融资；澜舟科技在2023年3月进行了Pre-A+轮数亿人民币融资，进行基础大模型训练；衔远科技于2023年3月完成数亿元人民币的天使轮融资，致力于链接消费者与商品。

相关研究

Aa Title	Time of Publication	Institute	Category	URL
GPT-4	2023.03.15	OpenAI	MLLM	https://cdn.openai.com/papers/gpt-4.pdf
Foundation Models for Decision Making	2023.03.07	Google	MLLM Plannning Review Robotics	https://arxiv.org/pdf/2303.04129.pdf
PaLM-E: An Embodied Multimodal Language Model	2023.03.06	Google	MLLM Plannning Robotics	https://arxiv.org/pdf/2303.03378.pdf
ChatGPT for Robotics: Design Principles and Model Abilities	2023.02.20	Microsoft	MLLM Robotics	https://www.microsoft.com/en-us/research/uploads/prod
SMART: SELF-SUPERVISED MULTI-TASK PRETRAINING WITH CONTROL TRANSFORMERS	2023.01.24	Microsoft	Control MLLM Robotics	https://arxiv.org/pdf/2301.09816.pdf
PERCEIVER-ACTOR: A Multi-Task Transformer for Robotic Manipulation	2023	NVIDIA	Control MLLM Robotics	https://proceedings.mlr.press/v205/shridhar23a/shridhar
RT-1: Robotics Transformer for Real-World Control at Scale	2022.12.13	Everyday Robots Google	Control MLLM Robotics	https://arxiv.org/pdf/2212.06817.pdf
A Generalist Agent	2022.11.11	DeepMind	MLLM Robotics	https://arxiv.org/pdf/2205.06175.pdf?fs=e&s=cl
VIMA: GENERAL ROBOT MANIPULATION WITH MULTIMODAL PROMPTS	2022.10.06	NVIDIA	Control MLLM Robotics	https://arxiv.org/pdf/2210.03094.pdf
Do As I Can, Not As I Say: Grounding Language in Robotic Affordances	2022.08.16	Everyday Robots Google	MLLM Plannning Robotics	https://arxiv.org/pdf/2204.01691.pdf

Aa Title	Time of Publication	Institute	Category	URL
LM-Nav: Robotic Navigation with Large Pre-Trained Models of Language, Vision, and Action	2022.07.26	Google	MLLM Navigation Robotics	https://arxiv.org/pdf/2207.04429.pdf
CLIPort: What and Where Pathways for Robotic Manipulation	2021	NVIDIA	Control MLLM Robotics	https://proceedings.mlr.press/v164/shridhar22a/shridhar
CLIP: Learning Transferable Visual Models From Natural Language Supervision	2021	OpenAI	CV MLLM	http://proceedings.mlr.press/v139/radford21a/radford21
PCT: Point cloud transformer	2021	Tsinghua	Point Cloud Transformer	https://link.springer.com/content/pdf/10.1007/s41095-02
Point Transformer	2021	HKU Oxford	Point Cloud Transformer	https://openaccess.thecvf.com/content/ICCV2021/paper
TransFusion: Robust LiDAR-Camera Fusion for 3D Object Detection with Transformers	2022	HKUST	LiDAR Point Cloud Transformer	https://openaccess.thecvf.com/content/CVPR2022/pape Camera_Fusion_for_3D_Object_Detection_With_Trans
LLaMA: Open and Efficient Foundation Language Models	2023.02.27	Meta	NLP Transformer	https://arxiv.org/pdf/2302.13971.pdf
PaLM: Scaling Language Modeling with Pathways	2022.10.05	Google	NLP Transformer	https://arxiv.org/pdf/2204.02311.pdf
Language Models are Few-Shot Learners (GPT-3)	2020	OpenAI	NLP Transformer	https://proceedings.neurips.cc/paper/2020/file/1457c0d6
BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding	2019	Google	NLP Transformer	https://arxiv.org/pdf/1810.04805.pdf&usg=ALkJrhhzxlCL
Attention Is All You Need	2017	Google	NLP Transformer	https://proceedings.neurips.cc/paper/2017/file/3f5ee243
ViT-22B: Scaling Vision Transformers to 22 Billion Parameters	2023.02.10	Google	CV Transformer	https://arxiv.org/pdf/2302.05442.pdf
MAE: Masked Autoencoders Are Scalable Vision Learners	2022	Meta	CV Transformer	https://openaccess.thecvf.com/content/CVPR2022/pape
Vision Transformer (ViT)	2021.06.03	Google	CV Transformer	https://arxiv.org/pdf/2010.11929.pdf